

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 19, 2003, 15:28:35 ; Search time 6310 Seconds

(without alignments)
4907.859 Million cell updates/sec

Title: US-09-494-297-2
Perfect score: 3945
Sequence: 1 MKKTRFPNKLNTLQFVLS.....IAGISLGIWGHIRIRKHD 757

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ .p2n.model -DEV=xlp
-O=/cgn2_1/USPRO.spool/US09494297/runat_13082003_122946_28092/app_query.fasta_1.903
-DB=GenEmbl -GPM=fastap -SUFFIX=rgc -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=bloum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09494297_CGN_1_1_4617_@runat_13082003_122946_28092 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEODDPRY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONCLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*
1: gb_da:*
2: gb_hcg:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_cm:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rdd:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3945	100.0	11289	1 AE006482	AE006482 Streptococcus
2	3943	99.9	2400	1 AB083107	AB083107 Streptococcus
3	2058.5	52.2	16207	1 AF447492	AF447492 Streptococcus
4	1965.5	49.8	10826	1 SF043937	SF043937 Streptococcus
5	1862	47.2	52900	1 AE014138	AE014138 Streptococcus
6	1862	47.2	311600	1 AP005141	AP005141 Streptococcus
7	1254.5	31.8	11420	1 AE009963	AE009963 Streptococcus
8	633.5	16.1	2845	1 SRRPRTNE	L10919 Streptococcus
9	611	15.5	2801	1 AF009908	AF009908 Streptococcus
10	603	15.3	2263	1 SPSEFB	X67947 S. pyogenes
11	523	13.3	2066	1 GG031115	U31115 Group G str
12	272.5	6.9	3221	1 AF009914	AF009914 Streptococcus
13	267.5	6.7	313	1 AF009913	AF009913 Streptococcus
14	241.5	6.1	326	1 AF009910	AF009910 Streptococcus
15	224.5	5.7	11264	1 AE009964	AE009964 Streptococcus
16	223	5.7	2202	1 AB084272	AB084272 Streptococcus
17	222.5	5.6	348	1 AF009920	AF009920 Streptococcus
18	211.5	5.4	278	1 AF009917	AF009917 Streptococcus
19	211.5	5.4	310	1 AF009911	AF009911 Streptococcus
20	210.5	5.3	309	1 AF009918	AF009918 Streptococcus
21	209.5	5.3	309	1 AF009915	AF009915 Streptococcus
22	207	5.2	301488	1 AE016950	AE016950 Enterococcus
23	205.5	5.2	282	1 AF009912	AF009912 Streptococcus
24	204.5	5.2	308	1 AF009919	AF009919 Streptococcus
25	203.5	5.2	54310	1 AP003515	AP003515 Clostridium
26	202.5	5.1	281	1 AF009909	AF009909 Streptococcus
27	200	5.1	301665	1 AE017009	AE017009 Bacillus
28	185	4.7	255	1 AF009916	AF009916 Streptococcus
29	180	4.6	2127	1 SEFNZEN	X99995 S. equi finz
30	178	4.5	2694	6 AR083961	AR083961 Sequence
31	178	4.5	7650	6 AR083960	AR083960 Sequence
32	178	4.5	8946	1 AF043131	AF043131 Moraxella
33	177	4.5	49617	6 AX067453	AX067453 Sequence
34	176	4.5	2718	1 AF043133	AF043133 Moraxella
35	175.5	4.4	1887	1 AF360373	AF360373 Streptococcus
36	171	4.3	301289	1 AE017000	AE017000 Bacillus
37	170.5	4.3	304708	1 AE017001	AE017001 Bacillus
38	168	4.3	300547	1 AE016954	AE016954 Enterococcus
39	164	4.2	3552	1 SF031980	U31980 Streptococcus
40	162.5	4.1	302470	1 AP000092	AP000092 Thermoplasma
41	162	4.1	5189	1 AY049089	AY049089 Streptococcus
42	161	4.1	4161	6 AX415264	AX415264 Sequence
43	161	4.1	4185	6 AX413693	AX413693 Sequence
44	161	4.1	4249	6 AR228645	AR228645 Sequence
45	161	4.1	4359	6 AR228644	AR228644 Sequence

RESULT 1

ALIGNMENTS

LOCUS	AE006482	11289 bp	DNA	linear	BCT 01-JUN-2001
DEFINITION	Streptococcus pyogenes M1 GAS strain SF370, section 11 of 167 of the complete genome.				
ACCESSION	AE006482 AE004092				
VERSION	AE006482.1 GI:13621422				
KEYWORDS					
SOURCE	Streptococcus pyogenes M1 GAS				
ORGANISM	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Streptococcus.				
REFERENCE	1 (bases 1 to 11289)				
AUTHORS	Perrett, J.J., McShan, W.M., Adjic, D., Savic, D., Savic, G., Lyon, K., Primeaux, C., Sezate, S., Surorov, A.N., Kenton, S., Lai, H., Lin, S., Qian, Y., Jia, H.G., Najjar, F.Z., Ren, Q., Zhu, H., Song, L., White, J., Yuan, X., Clifton, S.W., Roe, B.A. and McLaughlin, R.E.				
TITLE	Complete genome sequence of an M1 strain of Streptococcus pyogenes				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 98(48):14658-14663 (2001).				
MEDLINE	21192684				
PUBMED	11296296				
REFERENCE	2 (bases 1 to 11289)				
AUTHORS	Perrett, J.J., McShan, W.M., Adjic, D., Savic, D., Savic, G., Lyon, K., Primeaux, C., Sezate, S., Surorov, A.N., Kenton, S., Lai, H., Lin, S., Qian, Y., Jia, H.G., Najjar, F.Z., Ren, Q., Zhu, H., Song, L., White, J., Yuan, X., Clifton, S.W., Roe, B.A. and McLaughlin, R.E.				
TITLE	Direct Submission				
JOURNAL	Submitted (10-APR-2001) Department of Microbiology and Immunology, University of Oklahoma Health Sciences Center, 940 SL Young Blvd, Oklahoma City, OK 73104, USA				
FEATURES	Location/Qualifiers				
source	1. 11289				
gene	/organism="Streptococcus pyogenes M1 GAS"				
CDS	/mol_type="genomic DNA"				
gene	/strain="SF370"				
CDS	/serotype="M1"				
gene	/db_xref="taxon:160490"				
CDS	complement(1496..2137)				
gene	/gene="SpY0121"				
CDS	complement(1496..2137)				
gene	/EC_number="2.7.1.113"				
CDS	/note="Best Blastp hit = dbj BA016044.1 (AB030884)				
gene	similar to Lactobacillus acidophilus da/dg-kinase				
CDS	[Streptococcus zooepidemicus]				
gene	/codon_start=1				
CDS	/transl_table=11				
gene	/product="putative deoxyguanosine kinase/deoxyadenosine kinase(I) subunit"				
CDS	/protein_id="AAK33233.1"				
gene	/db_xref="GI:13621423"				
CDS	/translation="MLIVLAGTIGACKSSLAALGHEIGTDFYEAVDNNPVLAYYO				
gene	DPKKVAFILQIYFLNKRFSIKRAYOANNIILRSIFEDFLKLNKGNATKELD				
CDS	IYQELLANNIEELGEMPKRPDLITIDVSFDMLERIERGRSFEQVDCNPSELDIY				
gene	HOVGEPTWYEDYEVSPKMKIDGNSLDFVONPDATYLMKIDTKLKEHLHLL"				
CDS	complement(2157..3134)				
gene	/gene="SpY0122"				
CDS	complement(2157..3134)				
gene	/note="Best Blastp hit = gb AAC97155.1 (U49397) unknown				
CDS	[Streptococcus pyogenes]"				
gene	/codon_start=1				
CDS	/transl_table=11				
gene	/product="possible transcriptional regulator"				
CDS	/protein_id="AAK33234.1"				
gene	/db_xref="GI:13621424"				
CDS	/translation="MILNSSFRIIGVEIPHTVLAPMAGVTSAPRTTAKKEGAGIV				
gene	VMEIIEKGLVNNKTLHLHLIDENEPMSIOLFGDAGLKRADFIOTNPKAIIV				
CDS	DIMGCPVKNVYNEAGAKWLRDPIYHLYEVTVLIDPLTKMTGAGDSGLAVE				
gene	NALAASAGVSAALMAGRTREOMYTGCDHEETLARSKAITKIPITGNGDVSVOAK				
CDS	FMIEELGVAVMTGRAMNPNYLFQINHPFEGTGLPDLPAKAKIDIAADHLKRLIN				
gene	LKGETIAYREFRGLAPHYLRGTGAAKAVKAVSRAETLAEVEAIFETVR"				
CDS	complement(3121..3993)				
gene	/gene="SpY0123"				
CDS	complement(3121..3993)				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/gene="SpY0123"				
gene	complement(3121..3993)				
CDS	/				


```

QY      541 SerThrLeuAlaValAlaLysIleLeuValGluIurValIaGlnAspSerAsnProGln 560
Db      7516 AGTACTTACAGAGTTGCTAAATCTCTGTACAAATACGCTCAGATAGTAATCTCTCCACAG 7575
QY      561 LeuThrAspLeuAspPhePheIleProAsnAsnAsnLysTyrGlnSerLeuIleGlyThr 580
Db      7576 CTAACTGACCTTGATTCCTTATATCCGAATTAACAATTAATCAATCTCTATTTGGAAGT 7635
QY      581 GluThrPheProGluAspLeuValAspIleIleLeuGluMetGluAspLysGluValIle 600
Db      7636 CAGGCGATCCAGAAAGATTAGTTGATTAATTTGTTATGGAATGAAAAAGAAAGTTATA 7695
QY      601 ProValThrHisAsnLeuThrLeuArgLysThrValThrGluLeuAlaGlyAspArgThr 620
Db      7696 CCGTACTACTAATTTAACTTACAGAAAAACGGTGCAGCTGTTAGCTGGTGACAGAACT 7755
QY      621 LysAspPheHisPheGluIleGluLeuLysAsnAsnLysGlnIleLeuLeuSerGlnThr 640
Db      7756 AAGATTTCCATTTTGAATTTGAATTAATAAATGAAGCAAGAAATTCCTTCCAAACT 7815
QY      641 ValLysThrAspLysThrAsnLeuGluPheLysAspGlyLysAlaThrIleAsnLeuLys 660
Db      7816 GTTAAACAGATTAACAAACCTCGAATTTAAAGATGTAAGCAACCAATTAATTTAAAA 7875
QY      661 HisGlyLysLeuSerLeuThrLeuGlnGlyLeuProGluLysTyrSerTyrLeuValLysGlu 680
Db      7876 CATGGGAAAGTTAAACACTTCACAGGTTTACCAGAGTTATCTTACCTTGCTCAAGAA 7935
QY      681 ThrAspSerGluGlyTyrLysValLysValAsnSerGlnIleValAlaAsnAlaThrVal 700
Db      7936 ACAAGTTCTGAAAGCTTAAGGTTAAAGTTAATGTGCCAGAGCAAGCAATTCGTACAGTT 7995
QY      701 SerLysThrGlyIleThrSerAspGluThrLeuAlaPheGluAsnAsnLysGluProVal 720
Db      7996 TCAGAAAACAGGAATTAACAAGTATGATGACACCTCTCTTTGAAATTAATAAGACCTGTT 8055
QY      721 ValProThrGlyValAlaAspGlnLysIleAsnGlyTyrLeuAlaLeuIleAlaGly 740
Db      8056 GTTCCTACAGAGAGTTGATCAAAAGATCAATGGCATCTAGCTTGAATGATTAATGCGTGT 8115
QY      741 IleSerLeuGlyIleTrrPglyIleHisThrIleArgIleLeuArgLysHisAsp 757
Db      8116 ATCAGTTTGGGAGATCTGGGGAATTCACACGATTAAGATTAAGAAAAACATGAC 8166

RESULT 2
LOCUS   AB083107      2400 bp      DNA      linear      BCT, 21-JAN-2003
DEFINITION
  Streptococcus pyogenes cpal gene for collagen binding protein,
  complete cds.
ACCESSION
  AB083107
  AB083107.1 GI:23503477
SOURCE   Streptococcus pyogenes
  Streptococcus pyogenes
  Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
  Streptococcus.
REFERENCE
  1 Miyoshi-Akiyama,T., Wakisaka,N., Zhao,J. and Uchiyama,T.
  Characterization of collagen binding protein (cpa) of M1-type
  group A streptococcus
  Unpublished
  2 (bases 1 to 2400)
  Miyoshi-Akiyama,T., Wakisaka,N., Zhao,J. and Uchiyama,T.
  Direct Submission
  Submitted (03-APR-2002) Tohru Miyoshi-Akiyama, Tokyo Women's
  Medical University, Department of Microbiology and Immunology; 8-1,
  Kawada-cho, Shinjuku-ku, Tokyo 162-8666, Japan
  (E-mail: kfe01112@nifty.ne.jp; Tel: 81-3-3353-8111 (ex. 22713),
  Fax: 81-3-5269-7411)
  Location/Qualifiers
    1..2400
    /organism="Streptococcus pyogenes"
  FEATURES
    Source

```

```

/mol_type="genomic DNA"
/strain="291"
/serotype="M1"
/db_xref="taxon:1314"
/country="Japan"
/gene="cpa"
/gene="cpal"
94..2367
/gene="cpa1"
/codon_start=1
/transl_table=11
/product="collagen binding protein"
/protein_id="BAC20340.1"
/db_xref="GI:23503478"
/translation="MKTRPNNKLNLTNORVLSKSKRRTVTVGVFLMIPALYSM
VGAKTFGLVESTPNKINPDSSETRRWGIESTYVGHPIYKQFRAHDLRWLEBSR
SYGVYFENLKAFPLGSDSSVKKRWYKHDI SYKEDYALSPRIIDDELNOKLRAMY
NGRPONANGIMEGLIEPLNAIRVQEWVYSDNAPISNPDESKRESSENLVSTOLS
LMROALIKOLIDPMLATKMPKQVDDPOLSIFESEDKDKYKNGOMLISGLVPTKP
TPDDPMPNPQPTTSVLRKVAIGYSKLEAGATLOLDNVSFOARVSSNDIGE
RILSDGTVTLTRELNPAGYSIAEPIITFKVEAGKYVTIIDGRQIEMPNKEIVPSVE
AYNDPEEFSVLTTONYAKPYPAKNNKSSVYVCPNADLKSPPDSDEGCTMPDETT
GEVYTHIAGRDLPKTYVPRDPPDFLKHKKVIEKGRGQATIEVSGLETDLR
AAVQLAIYFTDSAEIDKDKLYHGFQNDSTLVAKLIVYADDSNPQLTIDF
FIPNNKYOGLIGTOWHPEDLVDIIRMEDEKVIPTVHNTLTKRTVYGLAGDKDFH
FEIELKNNKQELISQTVTKETDKTNLEFEDGKATINLKHESILTLQGLPEGYSYLVKED
SEGVKRVNSOEVAANATVSKTGTISDETLAEFNKKBPVPTGVQDKINGYLAIVLAG
ISLIGINGIHIRIRKHD"
BASE COUNT      873 a      364 c      463 g      700 t
ORIGIN

Alignment Scores:
Score:          9,11e-244      Length:      2400
Percent Similarity: 3941.00      Matches:      756
Best Local Similarity: 99.87%      Mismatches: 0
Query Match:      99.90%      Indels:      0
DB:                1      Gaps:      0

US-09-494-297-2 (1-757) x AB083107 (1-2400)
QY      1 MetLysLysThrArgPheProAsnLysLeuAsnThrLeuAsnThrGlnArgValLeuSer 20
Db      94 ATCAAAAAACAAGGTTTCCAAATTAAGCTTAATTAATCACTCAAAAGGCAATTAAGT 153
QY      21 LysAsnSerLysArgPheThrValThrLeuValGluValPheLeuMetIlePheAlaLeu 40
Db      154 AAAAAGCTCAAAAGCAATTAAGTCTGACCTTAAGTGGAGTCTTTTAATGATCTTCGCTTGG 213
QY      41 ValThrSerMetValGlyAlaLysThrValPheGlyLeuValGluSerSerThrProAsn 60
Db      214 GTAACTTCATGCTGGTGGTCTAAGACGCTTTTGGTTTAACTTAACAATCTCGACGCCAAGC 273
QY      61 AlaIleAsnProAspSerSerSerGluTyrArgTrrPtyrGlyTyrGluSerTyrValArg 80
Db      274 GCATTAATATCCAGATTCAGATTCGGAATACAGATGATGATGATGAATCTTATGTAAGA 333
QY      81 GlyHisProTrrTyrLysGlnPheArgValAlaHisAspLeuArgValAsnLeuGluGly 100
Db      334 GGGCATTCCCAATTAATAACAGTTTAAAGTACACAGATTTAAGGCTTAACCTTAAGAAGA 393
QY      101 SerArgSerTyrGlnValTrrCysPheAsnLeuLysLysAlaPheProLeuGluLysSerAsp 120
Db      394 AGTAGAAGTTATGATTAAGTTATTCCTTTAATTTAAAGAAAGCAATTTCTCTCGGATCAGAT 453
QY      121 SerSerValLysLysTrrPtyrLysLysHisAspGlyIleSerThrLysPheGluAspTyr 140
Db      454 AGTAGGTATAAAGGATTAATAAACAATGATGATGATCTCTCAAAATTTGAAGATTAAT 513
QY      141 AlaMetSerProArgIleThrGlyAspGluLeuAsnGlnLysLeuArgAlaValMetTyr 160
Db      514 GCGATTAAGCCCTTGAATTAAGGAGATGAGCTAAATCAGAAATTAAGAGCTTATGATAT 573

```

QY 161 AsnGlyHisProGlnAsnAlaAsnGlyIleMetGluGluGluProLeuAsnAlaIle 180
 |||||
 Db 574 AATGAGATCCACAAAATGCCAATGCTATTATGAGAGCGTTGGAACCTTGAATGCTATC 633
 QY 181 ArgValThrGlnGluAlaValIlePheTyrTyrSerAspAsnAlaProIleSerAsnProAsp 200
 |||||
 Db 634 AGACTTACACAAAGAGCGGATGCTACTATTCGATATATGCTCTATTTCTTAATCCAGAT 693
 QY 201 GluSerPheLysArgGluSerGluSerAsnLeuValSerThrSerGlnLeuSerLeuMet 220
 |||||
 Db 694 GAAAGTTTAAAGAGGAGTGAAGAACTGCTGTTAGCTCTCAATTTCTTTGATG 753
 QY 221 ArgGlnAlaLeuLysGlnLeuIleAspProAsnLeuAlaThrLysMetProLysGluVal 240
 |||||
 Db 754 CGTACAGCTTTGAAGCACTGATTCATCCGATTTGGCACTAAATAGCCAAACAGATT 813
 QY 241 ProAspPheGlnLeuSerIlePheGluSerGluAspLysGlyAspLysTyrAsnLys 260
 |||||
 Db 814 CCGGATGATTTTCAGCTAAGTATTTTGAGTCTGAGGCACAGGAGATTAATATATATAA 873
 QY 261 GlyTyrGlnAsnLeuSerGlyGlyLeuValProThrLysProProThrProGlyAsp 280
 |||||
 Db 874 GGATACCAAAATCTTTGAGTGGTGTGTTAGTCTCTAACAACCAACTCCAGGAGAC 933
 QY 281 ProProMetProProAsnGlnProGlnThrThrSerValLeuIleArgLysTyrAlaIle 300
 |||||
 Db 934 CCACCAATGCCCTCCAAATCAACCTCAACGACTTCAGTACTATTAGAAATATGCTATA 993
 QY 301 GlyAspTyrSerLysLeuGluGluGlyAlaThrLeuGlnLeuThrGlyAspAsnValAsn 320
 |||||
 Db 994 GGTGATTACTCTAAATGCTTGAAGGTGAGCAACATTACAGTTGACAGGGGATACGTGAT 1053
 QY 321 SerPheGlnAlaArgValPheSerSerAsnAspIleGlyGluArgIleGluLeuSerAsp 340
 |||||
 Db 1054 AGTTTCAAGGAGAGGAGGTTTGAAGTATGATTTGAGAGAAAGAAATTTGAATCTATCAGAT 1113
 QY 341 GlyThrTyrThrLeuThrGlnLeuAsnSerProAlaGlyTyrSerIleAlaGluProIle 360
 |||||
 Db 1114 GGAACTTATCTTAACTTGAATTAATTCACAGCTGCTTATGATATCGCAGACCAATC 1173
 QY 361 ThrPheLysValGluAlaGlyLysValTyrThrIleIleAspGlyLysGlnIleGluAsn 380
 |||||
 Db 1174 ACTTTTAAAGTTGAAGCTGGCAAGGTATCTATTAATGATGAAACACATGTAAGAAAT 1233
 QY 381 ProAsnLysGluIleValGluProTyrSerValGluAlaLysAsnAspPheGluLysPhe 400
 |||||
 Db 1234 CCCAATTAAGAGATGATGAGCCCTTACTAGTACGATACCATATTAATGATTTTGAAGAAATTT 1293
 QY 401 SerValLeuThrThrGlnAsnTyrAlaLysPheTyrTyrAlaLysAsnLysAsnGlySer 420
 |||||
 Db 1294 AGCGTTTAACTACACAAAATATGCAAAATTTTATATGCAAAAATTAATAATGGAAGT 1353
 QY 421 SerGlnValValTyrCysPheAsnAlaAspLeuLysSerProProAspSerGluAspGly 440
 |||||
 Db 1354 TCACAGGTTGCTCTATGCTTAAATGACAGATCTAAATCTCACACAGACTCGAAGATGCT 1413
 QY 441 GlyLysThrMetThrProAspPheThrThrGlyGluValLysTyrThrHisIleAlaGly 460
 |||||
 Db 1414 GGGGAAAACAATGACTCCAGACTTTTAAACAAGAGAACTAAATTAATCACTCATATATGCAAGT 1473
 QY 461 ArgAspLeuPheLysTyrThrValLysProArgAspThrAspProAspThrPheLeuLys 480
 |||||
 Db 1474 CGTGACCTCTTAAATATACGTGAACCAAGAGATACCCGACACTTCTTTTAAA 1533
 QY 481 HisIleLysLysValIleGluLysGlyTyrArgGluLysGlyGlnAlaIleGluTyrSer 500
 |||||
 Db 1534 CATATCAAAAAGAAATTAATGAGAGGTTTACAGGAAAAAGACAACTATATGAGTATAGT 1593
 QY 501 GlyLeuThrGlnThrGlnLeuArgAlaAlaThrGlnLeuAlaIleTyrTyrPheThrAsp 520
 |||||
 Db 1594 GGTCTAACTGAGACAAATTCGGCGGCTACAGTTAGCAAAATATATATTTACATGAT 1653
 QY 521 SerIleGluLeuAspLysAspLysLeuLysAspTyrHisGlyPheGlyAspMetAsnAsp 540

Db 1654 AGTCTGATTAAGTAAAGATTAACATAAAACATATCATGTTTGGAGACATGATGAT 1713
 QY 541 SerThrLeuAlaValAlaLysIleLeuValGluTyrAlaGlnAspSerAsnProGln 560
 |||||
 Db 1714 AGTACTTAAAGAGTGGCTAAATCTCTGTAGAAATACGCTCAAGATAGTAATCTCCACAG 1773
 QY 561 LeuThrAspLeuAspPhePheIleProAsnAsnAsnLysTyrGlnSerLeuIleGlyTyr 580
 |||||
 Db 1774 CTAAGTACCTTGAATTTCTTTATTCGAAATTAACATTAATTAATCAATCCCTTAATGGAAGT 1833
 QY 581 GlnThrPheProGluAspLeuValAspIleIleArgMetGluAspLysGluValIle 600
 |||||
 Db 1834 CAGTGGCATCCAGAAAGATTTAGTATTTTTCGTATGCAAGATTAATAAGAAAGATTATA 1893
 QY 601 ProValThrHisAsnLeuThrLeuArgLysThrValThrGlyLeuAlaGlyAspArgThr 620
 |||||
 Db 1894 CCTGTAACCTCATTAATTTAACTATGAGAAAAACGGTGAAGTGTGAGTGACAGAACT 1953
 QY 621 LysAspPheHisPheGluIleGluLeuLysAsnAsnLysGlnGluLeuLeuSerGlnThr 640
 |||||
 Db 1954 AAAGATTTCCATTTTGAATTTGAATTAATAATTAATGAAGCAAGAAATGCTTCCAAACT 2013
 QY 641 ValLysThrAspLysThrAsnLeuGluPheLysAspGlyLysAlaThrIleAsnLys 660
 |||||
 Db 2014 GTTAAACAGATTAACAAACCTCGAATTTAAAGATGTAAGCAACCATTAATTTAAA 2073
 QY 661 HisGlyGluSerLeuThrLeuGlnGlyLeuProGluGlyTyrSerTyrLeuValLysGlu 680
 |||||
 Db 2074 CATGGGGAAGTTTAAACCTTCAAGGTTTACCAAGGTTATTTCTTACCTGTCAAGAA 2133
 QY 681 ThrAspSerGluGlyTyrLysValLysValAsnSerGlnGluValAlaAsnAlaThrVal 700
 |||||
 Db 2134 ACAGATTTCTGAAGCTTAAAGTTAAAGTTAATGATGCAAGAGATGAGAAAGCTACAGTT 2193
 QY 701 SerLysThrGlyIleThrSerAspGluThrLeuAlaPheGluAsnAsnLysGluProVal 720
 |||||
 Db 2194 TCAAAACAGCAATTAACAGATGATGACAGACTTCCTTTGAAATTAATAAGACCTGTT 2253
 QY 721 ValProThrGlyValAspGlnLysIleAsnGlyTyrLeuAlaLeuIleAlaGly 740
 |||||
 Db 2254 GTTCTCTACAGAGATTTGATCAAAAGATCAATGCTATCTAGCTTGAATGATATGCTGCT 2313
 QY 741 IleSerLeuGlyIleThrGlyIleHisThrIleArgIleArgLysHisAsp 757
 |||||
 Db 2314 ATCAGTTTGGGATCTGGGATTCACACGATTAAGATTAAGAAACATGAC 2364
 RESULT 3
 AF447492
 LOCUS
 DEFINITION
 Streptococcus pyogenes RoFA (rofa), protein F (prtF1.12), Cpa (cpa.12), Lepa (lepa), EftLst.A (eftLst.A12), EftLst.B (eftLst.B), MsmRL (msmRL), and protein F2 (prtF2.12) genes, complete cds; and unknown gene.
 AF447492
 VERSION
 AF447492.1 GI:19224132
 KEYWORDS
 Streptococcus pyogenes
 Streptococcus pyogenes
 Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 Streptococcus.
 1 (bases 1 to 16207)
 AUTHORS
 Bessen, D.E. and Kalla, A.
 TITLE
 Genomic localization of a T serotype locus to a recombinatorial zone encoding extracellular matrix-binding proteins in Streptococcus pyogenes
 JOURNAL
 Infect. Immun. 70 (3), 1159-1167 (2002)
 MEDLINE
 21843101
 PUBMED
 11854196
 2 (bases 1 to 16207)
 REFERENCES
 Bessen, D.E. and Kalla, A.
 TITLE
 Direct Submission
 JOURNAL
 Submitted (12-NOV-2001) Epidemiology & Public Health, Yale

FEATURES University, 60 College Street, New Haven, CT 06520, USA
 source 1. 16207
 /location="Qualifiers"
 /organism="Streptococcus pyogenes"
 /mol_type="genomic DNA"
 /strain="A735"
 /serotype="M12"
 /db_xref="taxon:1314"
 /note="FCR region"
 complement(147..1466)
 /gene="rofa"
 /complement(147..1466)
 /gene="rofa"
 /note="transcriptional regulator"
 /codon_start=1
 /transl_table=11
 /product="Rofa"
 /protein_id="AAL86404.1"
 /db_xref="GI:19224133"
 /translation="MTIQRMISCOPTHPRSKETLYQLYASSNVQLALFLKNGSHS
 RPLTPFARSHPLNSGSAVMREALIPLNPFELSKNKTIVGEYRIRYLIALYKSF
 GIKYDLTGODKNITISPLSHSTHKTPMSESFPDILALSMKRHOSVITIPQ
 TRFDQKLFLYDSLKSSRDIIETVQDLSAGDLYLITANNFSALQMTPE
 EHRQCCQLFEENDTFRLNPIITLPLMLKEQASLVKALMEFSKSFENQHFPE
 TMLFSPYKGNOKLYTSKLIVEMMAKPGKRYLNHGHFLFCHYVBOILNIOP
 LVVVEAFNFINAHLTDSFPYSDKSIDFHSYLLDNVYOIPDLKPDLYVTHSOL
 LPVHEHETKGAIVAEISFDESILSIQELMTQYKEKKFQADLTKQLT"
 1872..3968
 /gene="prtfl.12"
 1872..3968
 /note="fibronectin-binding protein; protein Fl; SfbI"
 /gene="prtfl.12"
 /codon_start=1
 /transl_table=11
 /product="protein F"
 /protein_id="AAL86405.1"
 /db_xref="GI:19224134"
 /translation="MVSSYMFARGEKNNKMFLENKAGFLVHTKRRKRPVTLGVGF
 LLLACGALFGGVAADEKTYPNFKIDGSSAFKSTAAVRVLDENLKEIKLIL
 KEYOACFNLTKYFPRPYSTNNFKKIDGSSAFKSTAAVRVLDENLKEIKLIL
 NVTINCKSNANGFMGIEDLVNLTQNAIWTYSDAPLNDVKNKWESEVRNGEISE
 SOVTLNREALKILIDNLEATANKIPSGYRLNFKSENEIDONLSAEVDPDPKP
 GDTSEHNPTEPDLGPIPEDKRPDESEPALPLMPELDEGEVEVSESLPALP
 PLMPELDEGEVEVSESLPALPLMPELDEGEVEVSESLPALPLMPELDEE
 VPEVSESLPALPLMPELDEGEVEKPSVDLPITVPRYEENKNDOSLAGSGETE
 YITEVYGNQONPVYDIDKLPNNGEFGNAYETEDKEPEYLMGQSESEFTDITG
 MSGQITPOVEEDTKPEVLMGQSESEFTDITGMSGQITPOVEEDTKPEVLM
 GGQSESEFTDITGMSGQITPOVEEDTKPEVLMGQSESEFTDITGMSGQSF
 EYVTVEDTRPKLVFHFDDNNEPRKVEENREKPTKNIPILPATGDIEINVLAFGLILS
 VLSIFLKNKNKY"
 5326..7596
 /gene="cpa.12"
 5326..7596
 /gene="cpa.12"
 /note="putative collagen-binding protein"
 /codon_start=1
 /transl_table=11
 /product="Cpa"
 /protein_id="AAL86406.1"
 /db_xref="GI:19224135"
 /translation="MNNKILQKODAPRVSNRKPQTLTVGVFLMFLTLVSMKGA
 OSIIFGEKRIEESVPRKISPDAPYMGYDSDSHPYERKVAHDRLVNLNGSKS
 YOVYCPENINSHPNKRNAFSKQMKFVGDGTGVETNYAQTPIRGESLNNKLISITN
 AYPKANNGYMKDIEPLNATLVQAAWYISDSYGNKRTLMASELDGKIDPEQVYLM
 REAKSLISDDEETSKNLPQGSKLNTIVPOKSVONLSAEVPESPAPQSEEP
 PVQTKTSVILIRYVAGDYSLLEGATLRLTGEDLIDFQEKVQSGNGTEKLELSGT
 YLTETSSPDGYKIAEPIKFRVYNNKVFIVQKDSQVENVKAEVAVSEVYSDMD
 SHVINEFTPYGKFIYAKNKKDSQVYVCPADLHSSPESEGGGTIPDSTIMKEV
 KYTHGASDLFYALBRDTPEDFLKHIKVYTEKNGKDSYNGLTFTFOPPAAOL
 AITYETDSDTLTKTYNNGKGYHGFESDEKTLAKRLINAAQNSAPQLTNDPE
 VRNNSKYQSLIGTEHPDVLVIRHEDKQEVYTPHSHIVAKTYVGEIGDTKGFQ
 FELELKDGTGQIVNLTQNNODLVAKDKYSNKLHGDITRIEGLPTGYSTLAKETE
 AADYIVTQNKVSOEASASENVTADEKVTFENRKDLVPTPGITDGGTYLMLLIVP

gene
 CDS
 FGLVWFGRKGLKND"
 7589..8110
 /gene="lepa"
 7589..8110
 /gene="lepa"
 /note="putative signal peptidase I"
 /codon_start=1
 /transl_table=11
 /product="lepa"
 /protein_id="AAL86407.1"
 /db_xref="GI:19224136"
 /translation="MTNYLNRLENPLEFAFIRLVKISIIIGFLGYIEFOYFGWIV
 NTNQMSPAVSAGDGVLYRLTDYRHINDVYVVDVTLKGRISAQAGDVEFTQEG
 LLINCHPPEKEVPYLYTPHSSGNPFYKVPQTGYFILNDYRRELDRSYKALPNIQI
 KGRISFLRVRG"
 8132..9160
 /gene="eftLSL.A12"
 8132..9160
 /gene="eftLSL.A12"
 /codon_start=1
 /transl_table=11
 /product="eftLSL.A"
 /protein_id="AAL86408.1"
 /db_xref="GI:19224137"
 /translation="MKKNKLLATATLALGASLQNVKAETAGVSSGQDLTKKS
 ITNENDTLMPKTDYTFSVNPDASATGESNLPKIPGIAVNNODIKVYSMTDKTSG
 KKVQVVDVPMKTYFSPVGYRYVVPENKGTARGVYDTRKMLVDYVYGNNEKGLPEK
 YIVSKKGDSTAKERIQPNNSPFTSIIKEKYVNGCHDKAKFTPLTITQPNDEYAS
 SVYKIEENGQTDVKIGEAVKFTLNDQSVTLSKLPVGINTVVEEAEANQGGITTTAT
 LKDGKLSITYNLGOEHKTDKTADEIVVTVNNRDTQVPGVGLAPALVIAIGVI
 YIVRKKA"
 9176..9901
 /gene="eftLSL.B"
 9176..9901
 /gene="eftLSL.B"
 /codon_start=1
 /transl_table=11
 /product="eftLSL.B"
 /protein_id="AAL86409.1"
 /db_xref="GI:19224138"
 /translation="MMRTVQVYINNAIDTLLIFCLVLFPLAGFLGWSYHLXYQADA
 SNEKFKTAQOOQPEFDLALMEDVIGWLNIPGTHIDYLVQGTNLLEYINKAVGSA
 AMSGSLFETLRHNEDTDYSLIYGHMAGNMFGEIPKFLKDFENKINKAIIETKE
 RKRTIVTIRACLTQAFDQDLVNPNAITNQDQQLVDYISRSKQFQPKVYLKHH7KF
 VAFSTCENSTNRYVGTIDE"
 9936..10505
 /note="Orf2"
 /codon_start=1
 /transl_table=11
 /product="unknown"
 /protein_id="AAL86410.1"
 /db_xref="GI:19224139"
 /translation="MLFSVMIITMLAEVQYVLAKDSYVQTSISVENVLERGDSITPE
 SIALESIDAMKTIETITGASGKASFPLFTYVQGYTRYVYQKQBSQNNKDYQADTVF
 DVLVYTVYDEDTLVAKVISRRAGEKSAITFKRRLVKKIPPROPNIPKPLPLAG
 EVKSLGLISIVLGLVLLVYKLLSKSL"
 complement(10664..11869)
 /gene="msmRL"
 complement(10664..11869)
 /note="AraC-type DNA-binding domain-containing protein"
 /gene="msmRL"
 /codon_start=1
 /transl_table=11
 /product="msmRL"
 /protein_id="AAL86411.1"
 /db_xref="GI:19224140"
 /translation="WIFDLKEHVQTLHSLSPISYMSQKALIQYGGNDYLLCYQ
 FLKHLAIPQADAVIFEGILFPESEFIPFLCHYIIAIGFPEYSLNKDYQEOQLANNFL
 KHSSHKEKELLSYVALVHPPIPNVNRNLLIADAFPDQFETCOQITHOLOSQNL
 M7ADPDIIRLKHITSASSQDPVLELHINDVYKMGKQRLKQEIINPISLSTSS
 STSALAEKRNLTIVITRLLEFSVENDTVAAHYSIVAKYKMLNEASDILVYLIRG
 AAIIFHESELTKRSISDKROMNSVLAHYVDSHLYSKLVSDIAKRLYVSESLRSVFK
 KYSNWSIAHYIISTKIKEAQLLKRGIPVGEVAKSLYFVDTHFHKIRFKKYGISKD

gene
12254..15739
/gene="pRef2.12"
12254..15739
/gene="pRef2.12"
/note="fibronectin-binding protein; pfbp1"
/codon_start=1
/product="protein F2"
/protein_id="AAL6412.1"
/db_xref="GI:19224141"
/translation="MFOKNSYKILSFSLSTGFIQLGLLVFIIGLSGVSYGHAETRNGAN
KQSEFLIKVDONKPLPGATFSLTSKDGKSTVOTFTSNQKIVADONLQPGYTK
EETAPDGVKSTRTVYVYNGYKLYENYNEELISKASKRVSLSQLENKMSY
VSKYKTEVSSGADFTRNHAYTRKMSFLQKDKSETINGDFVLODLRLKPEYL
SODIPKTIYDSANPLAIGYHAENHOLIYFTDYIAGLDVOLSAELSLLEKPEYL
ENSTISNEKSTIGQELTYKGVNVLGNSTESNYITNLSNAGVSEIYVEPEGK
FMYVYVNPNTNIPATMNLMGFGARSKNSDLENDANTSALGEIOYVEPEGK
LPSYGVDTKRLRTDITAGLNGFQMTKORIDFGNIONKFAFIKYGKTDQSGK
PLVYOSNLASPRGASEYAAPTPGNCNYPFOEIALSPSKGSGSKSEPTKRSITYANL
KRVAQLRKKKSTDNVPLPEAFELRNSNGSKLEASNTQGEVHERDLTSGYIDL
ETKAPKIQOQTEKLAIVYDTKPAEEMVTGSPHSSVKYEANKEVITVNHKFTLIE
SGKRIWENDRPDPAKIOYLOLQNGKMPQIOEVKIDNDMSYHFDLPYDAKNDE
KYSEVEVNPDPGKYSLGNDIFNTEFEFEQNNNELEFGNAEIKGQSGSKTIDE
DLTSPFGKIKMKNDTAENRPOALIOVOLADGVAVEGOTKEISGNSMESEFNLIK
YNGTGNDIITYSVKEVYPTGVDTVYSAANDINTKREVTTOOGKPLETEPTEPESGAS
GGTTVEDSRVYDLSGLSSHQGSGDMTIEDSATIKRSKRIIDGKELGATMEIR
DSGKTIISTWISDQVDFYLMPEKTYFVEEAADPYEIALAIFTVNEOQOYVNEK
ATKQDTHIVWDAPKPKGSGOVNIDIEKLEPDEQHSSTIEIDSKSDLIIGQGS
VDPTEDTOSQMTGSHSGSTIEIDSKSDVLIIGQGOVETEDTQTMGYDSCKTE

Alignment Scores:

Pred. No.: 3.93e-122 Length: 16207
Score: 2058.50 Matches: 404
Percent Similarity: 71.24% Conservative: 126
Best Local Similarity: 54.30% Mismatches: 189
Query Match: 52.18% Indels: 25
DB: 1 Gaps: 10

US-09-494-297-2 (1-757) x AF447492 (1-16207)

QY 17 ArgValLeuSerLysAsnSerLysArgPheThrValThrLeuValGlyValPheLeuMet 36
Db 5365 CGGGTATCAACAAGAACCAAAACAATTACTGTCACCTTAAAGGGGATTTTATG 5424
QY 37 IlePheAlaLeuValThrSerMetValGlyValAlaLysThrValPheGly----- 52
Db 5425 TTTTGAACCTTGTAAGTTCATGAGAGGTCCTCAAGCATATTTGGAGAGAAAGAGA 5484
QY 53 LeuValGluSerSerThrProAsnAlaIleAsnProAspSerSerSerGlyTyrArgTyr 72
Db 5485 ATTAAAGACGTCAGTCTTCCATAAATAAAGTCCAGATGATGCC-----TACCCTTGG 5538
QY 73 TyrGlyTyrGluSerTyrValArgGlyHisProTyrTyrLysGlnPheArgValAlaHis 92
Db 5539 TATGCGTATGATCATATGACTCATGCAATCCTTACATGAAACGTTTAAAGTAGCAAT 5598
QY 93 AspleuArgValAnleuGluGlySerArgSerTyrGlnValTyrCysPheAsnLeuLys 112
Db 5599 GATTAAAGCGTTAATTAAATGAAGTAAGCTAACCAAGTATATTCGTTTAAATATCAAT 5658
QY 113 LysAlaPheProLeuGlySerAspSerSerValLysLysTyrPyrLysLysHisAspGly 132
Db 5659 TCTCATTTATCCGAATACAAAAAATGCTTTTCTTAACAATAGTTTAAAGAGATTGATGG 5718
QY 133 IleSerThrLysPheGluAspTyrAlaMetSerProArgIleThrGlyAspGluLeuAsn 152
Db 5719 ACAGGTGATGTCACAAATATTATGTCAGACACCTTAAGATTCTGGAGATTCATTGAT 5778
QY 153 GlnLysLeuArgAlaValMetTyrAsnGlyHisProGlnAsnAlaAsnGlyIleMetCln 172
Db 5779 AATTAACGTTTAAATGATATGATACACGCTATATCTTAAATAATGCTAATAGCTATATGAT 5838

QY 173 GlyLeuGluProLeuAsnAlaIleArgValThrGlnGluAlaValThrTyrTyrSerAsp 192
Db 5839 AAGATAGAACCATTAATGCTATTTTAACTAAGCTCAACAGCGTTTGGTACTATTTCTAC 5898
QY 193 AsnAlaProIleSerAsnProAspGluSerPheLysArgGluSerGluSerAsnLeuVal 212
Db 5899 AGTTCT---TATGGTAATATTAATAAACGTTATGGCATCTGACCTTAAGAGGAAATA 5955
QY 213 SerThrSerGlnLeuSerLeuMetArgGlnAlaLeuLysGlnLeuIleAspProAsnLeu 232
Db 5956 GATTTTGAACAAGTAATTAATGGCTGAACCTTACTCAAAACCTTAAATGATGATGATTA 6015
QY 233 AlaThrLysMetProLysGlnValProAspAspPheGlnLeuSerIlePheGluSerCln 252
Db 6016 GAAGAACATCTAAATAATACCTACCTCAAGCATCTAAGCATTAATGATTTTGTCCGCA 6075
QY 253 AsplysGlyAspLysTyrAsnLysGlyTyrGlnAsnLeuLeuSerGlyLysLeuValPro 272
Db 6076 GAT-----AAATCTGTCAAAATTTATTAAGTCAGACGATGCTGCT 6117
QY 273 ThrLysProProThrProGlyAspProMetProProAsnGlnProGlnThrThrSer 292
Db 6118 GAATCCCTCCGCGACAGTCAGTCTCAAGAACCGCAGCAAAACAAACAAACATCA 6177
QY 293 ValLeuIleArgLysTyrAlaIleGlyAspTyrSerLysLeuLeuGluGlyAlaThrLeu 312
Db 6178 GTCAATTATCAAAATATGCGAAGGTACTACTTAACTTAACTTAAAGGAGCAACTTGG 6237
QY 313 GlnLeuThrGlyAspAsnValAsnSerPheGlnAlaArgValPheSerSerAspIle 332
Db 6238 GCTTAAACAGGGAAGATATCTGATTTTCAAGAAAAGCTTCCAAAGTAATGGAACA 6297
QY 333 GlyAlaTyrGlyLeuLeuSerAspGlyThrTyrThrLeuThrGluLeuAsnSerProAla 352
Db 6298 GGAGAAAAGATTGATTAATCAATGGGACTTAACTTAAACGAACATCACTCCAGAT 6357
QY 353 GlyTyrSerIleAlaGluProIleThrPheLysValGluAlaGlyLysValTyrThrIle 372
Db 6358 GCATTAATAATTGCGGAGCCCATTAAGTTTAAAGTAGTAATTAATAAAGTATTTATGCTC 6417
QY 373 --IleAspGlyLysGlnIleGluAsnProAsnLysGluIleValGluProTyrSerVal 391
Db 6418 CAAAAGATGGTCTCAAGTGAAGTAATCAAAACAAAGATGACGACCATACTCACTG 6477
QY 392 GluAlaTyrAsnAspPheGluGluPheSerValLeuThrThrGlnAsn-----Tyr 408
Db 6478 GAACGTAACAGCATATGCAAGATAGTACTATTAATTAATCCAGAAACCTTCACTCTTAT 6537
QY 409 AlaLysPheTyrTyrAlaLysAsnLysAsnGlySerSerGlnValTyrCysPheAsn 428
Db 6538 GCGAATTTTATTCGCTAAATAATAGATTAAGATTAAGATGTTGTTACTGTTTAT 6597
QY 429 AlaAspLeuLysSerProProAspSerSerGluAspGlyLysThrMetThrProAspPhe 448
Db 6598 GCTGATTACACTCTCCACCTGAATCAAGAGATGGGAGGAGACATATGATCAGATATT 6657
QY 449 ThrThr---GlyGluValLysTyrThrHisIleAlaGlyArgAspLeuPheLysTyrThr 467
Db 6658 AGTACGAGAAAGATCAAGTACACATGACGATCGGAGTAGTATGTTTAAATATACCG 6717
QY 468 ValLysProArgAspThrAspProAspPheThrPheLeuLysHisIleLysValIleGlu 487
Db 6718 CTAAGACCGAGAGATCAAAATCCAGAACACTTCTTAAAGCACTTAAATAAGTAATGAA 6777
QY 488 LysGlyTyrArgGluLysGlyGlnAlaIleGluTyrSerGlyLeuThrGluThrGlnLeu 507
Db 6778 AAAGGCTACAAATAAAGATGATAGC-----TATATGATTAACAGAAACACAGATT 6831
QY 508 ArgAlaIaThrGlnLeuAlaIleTyrTyrPheThrAspSerAlaGluLeuAspLysAsp 527
Db 6832 CGCGCGGCTACTACGCTTGTATCTATTAATCTTACAGACACCTGACTGATTAACCTTA 6891
QY 528 LysLeu-----LysAspTyrHisGlyLysPheGlyAspMetAsnAspSerThrLeu 543


```
/product="Cpa"
/protein_id="AAC97148.1"
/db_xref="GI:4033711"
/translation="MQRKDTNGSANKRRQTTIGLKVLFVVALIGVFSTRAP
GAEDSVPNRQSSVODYPWYDGYPKGYSPDKTHTNKLKVNLESGDKDOACFNL
TKHPPSKDSVSRSOMYKLEGTENFIKLADKPRIEDGLOONLRIILNGYPPNNG
IMKIDPLNALITONAIWYDSAOINDESEFKTEARSNGINDOOLGIMRKALKEID
PNLSKSNKTPSCYRLNWFESHDKPRONLSAEVPTPPKPEEPKTEKTSVIL
RKVAGDSKLEAGATLKLISQLESGFQEKDQSNLSGETVELPGTYTLTETSSPDG
KIAEPKIFERNKVVFIQKDSQVENEKPEVAPYVEAYANDMEDEVLSGFTPYK
FYATNMKDSQVYVCFNADLHSPDSQSEITINPDSTMKEVKYHTTAGSADLEKVA
LRPDNTEDEFLKIKKYEKKKGSYNGLETOPRAFATOLAIFYETGSDLEKTL
KTYNNGDGHGEFSEMDKTLAVTEKLITYAQNGSAPOLINDFPVNNKXDSILGTE
CHPDDLVDTVMEDKKOEVLPVTHSHSTPYKKTVCGLGDKTGFQPELEKKTQPIV
NTLKTNNODLVAKDKTSFNLKHGDTITIEGLPGYSTTLKEAKKDTIVYVKNVSO
EAQSVGKDITEDKKVTFENRKDLVPPETGLTDGAIYIMLLLVPLGLLWLFGRGLK
ND"
complement(5972..5977)
/gene="cpa"
/note="promoter 1"
complement(5997..6102)
/gene="cpa"
/note="promoter 1"
complement(6138..6148)
/gene="cpa"
/note="extended promoter 2"
complement(6161..6166)
/gene="cpa"
/note="promoter 2"
6231..7898
/gene="nra"
6231..6241
/gene="nra"
/note="nra"
/note="extended promoter 2"
6297..6302
/gene="nra"
/note="nra"
6315..6325
/gene="nra"
/note="extended promoter 1"
6363..7898
/note="extended promoter 1"
/gene="nra"
/note="negative regulator of group A Streptococci"
/codon_start=1
/transl_table=1
/product="Nra"
/protein_id="AAC97147.1"
/db_xref="GI:4028948"
/translation="MPYKKKKDSFIVETYLEOSTRKSSELYLLFKSPITIFSVAK
OTGLAVOLAKYCKELDDPFGNNNDITKKKIKCCPVKPKFETLHOLYDTSTILKL
LVFPIKNGTSSQPLIKFSKRYLSSSSVYRRESLILKLRFGKLVSKNTVGEYRI
RYLLIAMYSKGIYIPLDHDNLIYRFLSOSATNLSTPMLEPFSFYMLLALSW
KRHOFAVSIPTORIFROLKLFIDYCLRSSROYENAFSLFSGGDDYFLIYITW
NSFASLOWTPOHIECTCHIEKNDTEFLLEPLIKRLPOLNHSKODLIKAMTESK
FLNLOHFEVIEIPESFLPYTNGSNLYALKNTYINQMLAOLPKRHLNEKHLQFCSH
IEQLKKNOPALYVLLISSNFINAKLIDTIPRYSDKGIFHTSYLLKRDITQIPSL
KPDLYITHSRILPVPKNDLVKGVTAERSFDPKPYSIASIONLIYOLKDKRYODFLNE
OLQ"
8006..8016
/note="ORF 5; extended -10 signal"
8024..8857
/note="ORF 5"
/codon_start=1
/transl_table=1
/product="unknown"
/protein_id="AAC97154.1"
/db_xref="GI:4033717"
/translation="MDKIKSIAGSGAFRAYLDSTETVALAQEKHNTLSSITYALGR
TLIANQILANQKDSKITVAVIGDSSGHIISVADTGHVGYIQNGVIDIKRTENG
EVLVGFPMNGHFVTIIDYGNPYSTPTITIGEDPFIYTLIESQTPSALGNL
LIDENDKVKVAGGVAVLPEASEEIRYKRLQEMPAIHLILASKHNVALLAEAY
GDEPYKRLSEPLSFQCCCSREFFEAALMTLPKADLQAMIDEDKGAELIVQCFCKLPL
I"
```

```
CDS
8884..9861
/note="ORF 6; similar to Azospirillum brasilense nifH3
gene"
/codon_start=1
/transl_table=1
/product="unknown"
/protein_id="AAC97155.1"
/db_xref="GI:4033718"
/translation="MIKLNSSFRIGDVEIPIRVLAPMAGVYNSAFRTIAEPGAGLV
VMEIISEGLLYNNEKTLHMLHIDENEPMSIQLEGGAEGLKRAADFIQNTKADIY
DINMGCPVAKVKNKEAGAKMLRDPKITYHIVEYVSADILPYVMRTGMADSLAVE
NALAAESGVALMAGHTRTROMYGTGDDHEFLKVAISAIRKIPKIGNDVRSODAK
EMIEIGDAVMIEGMAANNNYLTQIHFEETGCELPDLPPAKKIDJAKHKLRLIN
LKGETIARBEREGLAIPHRLRTAGCAAKRGAVSRNETIAEVALFEYR"
9881..10522
/note="similar to Lactobacillus acidophilus da/dG-kinase"
/codon_start=1
/transl_table=1
/product="unknown"
/protein_id="AAC97156.1"
/db_xref="GI:4033719"
/translation="MLIVLAGTIGAKSSLAALGELGTDVFEAVDNNPVLDFYQ
DPKRYAFLOIYFLNKRREKSIKEAYQADNNLIDRSIFEDFLKLYNKNGVVTELD
IYQELLANMLELEBGMKRPDLIYIDVSPDKMERIEKRGSRPEQVQDNPSEQY
HOVGEYPTWYEDKVPSPKMKIDGSLDFVQNPDLATVLMIDTKLKEHL"
10598..10681
/note="putative transcription terminator"
/rpl_type=inverted
complement(10782..>10826)
/gene="ssba"

repeat_region
gene
Alignment Scores:
Pred. No.: 2.3e-116 Length: 10826
Score: 1965.50 Matches: 409
Percent Similarity: 70.00% Conservative: 116
Best Local Similarity: 54.53% Mismatches: 192
Query Match: 49.82% Indels: 34
Dbs: 1 Gaps: 16

US-09-494-297-2 (1-757) x SP049397 (1-10826)
QY 20 SerLySasSerLyArg-----PheThrValThrLeuValGlyValPheLeuMetIle 37
DB 5903 AGCGCTAACACAAACGACGACAAACGACGATTCGTAAGTATTTTTCGCGTTT 5844
QY 38 PheAlaLeuValThrSerMetValGlyValThrValPheGlyLeuValGluSerSer 57
DB 5843 GTAGCTCGATGAGTAATAGTAGCTTTCTATCAAGAGCGTTCGGAGCGAAGAACATCA 5784
QY 58 ThrProAlaAlaIleAsnProAspSerSerSerGluTyraGtTPTyGlyTyGluSer 77
DB 5783 GTACCAAAAT-----AGACAAAGCTCAATTCAGATTCGTCGTATGCGCATGATTCCT 5730
QY 78 TyrValAlaGlyIleAsnProTyThrTyTrGlnPheArgValAlaIleAsnLeuArgValAsn 97
DB 5729 TATCCTAAAGGCTACCCAGACATATGTCGTTAAAGACTTACCATTAATTTAAAGTAAT 5670
QY 98 LeuGluGlySerArgSerTyTrGlnValTyrcysPheAsnLeuIleValAlaPheProLeu 117
DB 5669 TTAGAGGAGAGTAAGATATTCACACATACGCTTATTTTAAACAAACATTTTCATCC 5610
QY 118 GlySerAspSerSerValIleValTyTrTyTrIleValIleAsnGlyIleSerThrIleSph 137
DB 5609 AAGTCAGATAGTGTAGATCACAAATGTTAAAAAATGTAAGGAACTAAAGAAACATTT 5550
QY 138 GluAspTyrrAlaMetSerProArgIleThrGlyAspGluLeuAsnGlnIleLeuArgAla 157
DB 5549 ATCAAGTTAGCAGATTAACCAAGATTAAGACGACAGCTTACCAACAAATATATTGAGG 5490
QY 158 ValMetTyraAsnGlyIleProGlnAlaAsnGlyIleMetGluGlyLeuGluProLeu 177
DB 5489 ATTCTCTATATATGATATCTATATATATGCTAATATGCTAATATGAAGGATGATTCCTGA 5430
```


TITLE
JOURNAL

Direct Submission
Submitted (14-JUN-2002) Laboratory of Human-Bacterial Pathogenesis,
Rocky Mountain Laboratories, NIAID, NIH, 903 South Fourth St.,
Hamilton, MT 59840, USA

FEATURES
source

1. 52900
Location/Qualifiers

/organism="Streptococcus pyogenes MGAS315"
/mol_type="genomic DNA"

/strain="MGAS315"

/serotype="M3"

/db_xref="taxon:198466"

/note="group: A"

77. 1015
/gene="comYA"

/note="synonym: Spym3_0078"

77. 1015
/gene="comYA"

/note="best non-GAS blastp hit: gb|AAC45310.1| (U81957)

putative ABC transporter subunit ComYA [Streptococcus

gordonii]"

/codon_start=1

/transl_table=1

/product="putative ABC transporter subunit ComYA"

/protein_id="AAW78685.1"

/db_xref="GI:21903796"

/translation="MNOALAKAIIKAKAONOADIYILPRADQYDIFRIGDERLVD

YVQSDRAAPLISHKRFVAGMIVGERKRCQVSCDYKLSKQSLRLSSVGDYRQGS

LVIRLHHQNSVHYWFDGLKRVANQVGGRLYLFAGVSGKTTLMQLISSYHOEA

QVISEDPEVEIKNHQILQLOVNDIGMTYDNLIKLSRHRDILVIGETIRSORARAV

IRDSLGTGMVFSTVHAKSISGVARLELGTVKELSNCLALIAVORLLNGCALIDST

QNEFEYSSNMNQIIDQLLEAGHLNPKQAKLEKII"

951. 1985
/gene="comTB"

/note="synonym: Spym3_0079"

951. 1985
/gene="comTB"

/note="best non-GAS blastp hit: gb|AAC45311.1| (U81957)

putative ABC transporter subunit ComTB [Streptococcus

gordonii]"

/codon_start=1

/transl_table=1

/product="putative ABC transporter subunit ComTB"

/protein_id="AAW78686.1"

/db_xref="GI:21903797"

/translation="MISFLRQDISIPSKLSKSKHQHKEFIOLLANLSTGFSFAE

VIAFLKRSQILQDLYVMESLSLKGGLADMLSGLSDAILTOISLADHGNIEY

LVAIQHYLNQARIRKRVETVYPLILFLFVMMGLRRLVLPQLETOHQIYFLN

HPAPFIFGCSGLILFGMWLRMSOSRLKYSRLSRYPLGKLLKOYLSYAREN

GTLLGGIDMTIIDIMAIKESLSMKELAEIDIRMSLLEGSAFHIKVATYPPFKKELSL

MIYGEIKSKIGALEIYAGSEMFQPSQLYQVLIQIOPALFLVVAVTIMYAAIIL

PIYOMGIF"

1987. 2313
/gene="comYC"

/note="synonym: Spym3_0080"

1987. 2313
/gene="comYC"

/note="best non-GAS blastp hit: gb|AAC45312.1| (U81957)

ComYC [Streptococcus gordonii]"

/codon_start=1

/transl_table=1

/product="putative competence protein"

/protein_id="AAW78687.1"

/db_xref="GI:21903798"

/translation="MINQWNLRRHKKLKGFTLLEMLLIVIVISVLMLEFNLSKQKD

RVTETGNAAVKLVENQAELEYLSQSKPSLSQKADGSIITEKQKAYADYDKHKN

KARLSN"

2273. 2716
/gene="Spym3_0081"

2273. 2716
/gene="Spym3_0081"

/note="best non-GAS blastp hit: gb|AAL00663.1| (AE008550)

Competence protein [Streptococcus pneumoniae R6]"

/codon_start=1

/transl_table=1

/product="putative competence protein"

/protein_id="AAW78688.1"

/db_xref="GI:21903799"

/translation="MTNIMKRPVLAIAKAFLLLETLISLVMSFFIILGSLVPVTKSYQ

KVEELHFFSHFEHLRYHQOKLAIILOQKORVLIDISTKLVTEGNSLTVPKSITVNHPR

LVIDOMGNHSLAKIIFDMTRRFKRYQGLSSGNVQKFSQSLHSP"

2673. 2957
/gene="Spym3_0082"

2673. 2957
/gene="Spym3_0082"

/note="best non-GAS blastp hit: gb|AAK06217.1| (AE006440.3)

(AE006440) hypothetical protein [lactococcus lactis subsp.

lactis]"

/codon_start=1

/transl_table=1

/product="conserved hypothetical protein"

/protein_id="AAW78692.1"

/db_xref="GI:21903803"

/translation="MTPEKIEAVQQLLENCOLIEINDLKTTHYDAIVEONSFYGAEG

ASPYAONSDKLKALCKTEEMKRAYOGLFKAOTBQLOAHQFPTPAIGFILLYL

EOLSKDSTLEVLITGSGTGNLAOTLLNNTSKSLDYVGIETLDDLLIDLSASIAELMDS

AHFTQEDAVRPQLLAKESDIYISDLPVGYTPNDIDAKRKVASSKHTYAHHLMEQSL

KYLRKDGFAITFLAPVNLTSPOSOLKQWLVDAQVVALINLPDSIFGHPNSAKSIIY

LQKQIDHPEFETVYPIRDLKLAENIHDEMFENFKWKLISNVN"

4804. 6000
/gene="ackA"

/note="synonym: Spym3_0086"

/transl_table=1

/product="putative competence protein"

/protein_id="AAW78689.1"

/db_xref="GI:21903801"

/translation="MSKOLSNKAPFTLEALIALLVISGSLVYQGLNQTLLKRSHYL

ARHODMWLFRSHORELSGARPYKVDNLTVEKCKKYLAFQCFKSHDRKASNSG

KGYQMLFGISRSHIHESQICITLKKSGLERFYAFOD"

3368. 3694
/gene="Spym3_0084"

3368. 3694
/gene="Spym3_0084"

/note="best blastp hit: gb|AAK33225.1| (AE006481)

hypothetical protein [Streptococcus pyogenes M1 GAS]"

/codon_start=1

/transl_table=1

/product="hypothetical protein"

/protein_id="AAW78691.1"

/db_xref="GI:21903802"

/translation="MPRTNLRKAGILLYAMFMASIFLLVLOYIYSQVYALKREYQAOI

DYKARLAEIVYODHRYKASNPVFFKGQVYCRERRKMMIYKLDQROYOFEYIK

PKOYSO"

3792. 4745
/gene="Spym3_0085"

3792. 4745
/gene="Spym3_0085"

/note="best non-GAS blastp hit: gb|AAK06112.1| (AE006431.4)

(AE006431) hypothetical protein [lactococcus lactis subsp.

lactis]"

/codon_start=1

/transl_table=1

/product="conserved hypothetical protein"

/protein_id="AAW78692.1"

/db_xref="GI:21903803"

/translation="MTPEKIEAVQQLLENCOLIEINDLKTTHYDAIVEONSFYGAEG

ASPYAONSDKLKALCKTEEMKRAYOGLFKAOTBQLOAHQFPTPAIGFILLYL

EOLSKDSTLEVLITGSGTGNLAOTLLNNTSKSLDYVGIETLDDLLIDLSASIAELMDS

AHFTQEDAVRPQLLAKESDIYISDLPVGYTPNDIDAKRKVASSKHTYAHHLMEQSL

KYLRKDGFAITFLAPVNLTSPOSOLKQWLVDAQVVALINLPDSIFGHPNSAKSIIY

LQKQIDHPEFETVYPIRDLKLAENIHDEMFENFKWKLISNVN"

4804. 6000
/gene="ackA"

/note="synonym: Spym3_0086"

Db	17112	ACACAGGCTATCATGTTTGGATTAACCTGATCATCAACATTAGCTGATGATCATGAG	17171
Qy	549	LeuValGluTyrAlaGlnAspSerAsnProGlnLeuThrAspLeuAspPheIle	568
Db	17172	TTGATCTACCTATGCTGATGATGTCACCTTACCAATGATGATCAAAATTAAGATTCTTGT	17231
Qy	569	ProAsnAsnAsnLysTyrGlnSerLeuIleGlyThrClnTrpHisProGluAspLeuVal	588
Db	17232	CCPAAATAGCAGTATGATACAGGACCTATTGGAACGAGATATCAATGAAATGAAATGATT	17291
Qy	589	AspIleLeuArgMetGluAspLysLys---GluValIleProValThrHisAsnLeuThr	607
Db	17292	GATCTTATTCTTATGAGAGATAAACAGGCTCCTATTCTTCTATTACTCACAATTACT	17351
Qy	608	LeuArgLysThrValThrGlyLeuAlaGlyAspArgThrLysAspPheHisPheGluIle	627
Db	17352	ATTTCTAAACCTGTTACTGAGACTATTGACAGATTAAGAAAAAGAAATTAATTAATTAATA	17411
Qy	628	GluLeuLysAsnAsnLysGlnGluLeuLeuSerGlnThrValLysThrAspLysThrAsn	647
Db	17412	CATTTAAATCTCTGATGACAGCAAGCTATAGTGAACATATCCGACAACTCGAGAA	17471
Qy	648	LeuGluPheLysAspGlyLysAlaThrIleAsnLeuLysHisGlyLysSerLeuThrLeu	667
Db	17472	CTCAGACTTACAGATGAAAAAGCTACTTACATTAAGATGAGATGAGATCATTTGATTGT	17531
Qy	668	GlnGlyLeuProGlnGlyTyrSerTyrLeuValLysGluThrAspSerGlnGlyTyrLys	687
Db	17532	GAGGGGCTACCTCAGCTACTCTTATGAATACAGAAAGGGGCTGCTCAGATTATGAG	17591
Qy	688	ValLysValAsnSerGlnGluValAlaAsnAlaThrValSerLysThrGlyIleThrSer	707
Db	17592	GTAAGTTTATGAGAAAAATGACACGAGATGAAAAACGAGAAAGCTCAGTTAAGAA	17651
Qy	708	AspGluThrLeuAlaPheGluAsnAsnLysGluProValValProThrGlyValAspGln	727
Db	17652	GATGAGCTGATCTTTTGAACCCGAAAGATCTTCTCCACCACTGCTTT-GACAAC	17710
Qy	728	LysIleAsnGlyTyrLeuAlaLeuIleValIleAlaGlyIleSerLeuGlyIleTrp---	746
Db	17711	AGATGG---GGCTATCATCTCTTGGTGTATTACTTGG-TTCCATTGGGTTATTGCTTT	17766
Qy	747	GlyIleHisThrIleArgGlyIleArgLysHisAsp	757
Db	17767	GGCTATTGTCCTGAAAGGACCTAATAAATGAC	17799
RESULT 6	AP005141	311600 bp DNA linear BCT 13-JUN-2003	
LOCUS	AP005141	BA000034	
DEFINITION	Streptococcus pyogenes SSI-1 DNA, complete genome, section 1/6.		
ACCESSION	AP005141.1	GI:28810157	
VERSION	AP005141.1	GI:28810157	
KEYWORDS	Streptococcus pyogenes SSI-1		
SOURCE	Streptococcus pyogenes SSI-1		
ORGANISM	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Streptococcus.		
REFERENCE	1		
AUTHORS	Nakagawa, I., Kurokawa, K., Yamashita, A., Nakata, M., Tomiyasu, Y., Okashashi, N., Kawabata, S., Yamazaki, K., Shiba, T., Yasunaga, T., Hayashi, H., Hattori, M. and Hamada, S.		
TITLE	Genome Sequence of an M3 Strain of Streptococcus pyogenes Reveals a Large-Scale Genomic Rearrangement in Invasive Strains and New Insights into Phage Evolution		
JOURNAL	Genome Res. 13 (6), 1042-1055 (2003)		
MEDLINE	22683278		
REFERENCE	2 (bases 1 to 311600)		
AUTHORS	Yamashita, A., Nakagawa, I., Kurokawa, K., Nakata, M., Tomiyasu, Y., Yamazaki, K., Okashashi, N., Kawabata, S., Yasunaga, T., Hattori, M., Hayashi, H. and Hamada, S.		
TITLE	Direct Submission		
JOURNAL	Submitted (01-MAY-2002) Ken Kurokawa, Osaka University, Genome		

FEATURES	source	COMMENT
gene	gene	Information Research Center: 3-1, Yamadaoka, Suita, Osaka 565-0871, Japan (E-mail: kengen-info.osaka-u.ac.jp, URL: http://www.gen-info.osaka-u.ac.jp/, Tel: 81-6-6879-8365, Fax: 81-6-6879-2047)
CDS	CDS	genome project This clone was isolated from a patient presenting with toxic shock like syndrome.
Location/Qualifiers	Location/Qualifiers	
1..311600	1..311600	
/organism="Streptococcus pyogenes SSI-1"	/organism="Streptococcus pyogenes SSI-1"	
/mol_type="genomic DNA"	/mol_type="genomic DNA"	
/strain="SSI-1"	/strain="SSI-1"	
/db_xref="taxon:193567"	/db_xref="taxon:193567"	
/gene="SpS0001"	/gene="SpS0001"	
232..1587	232..1587	
/gene="SpS0001"	/gene="SpS0001"	
232..1587	232..1587	
/note="similar to GB:AA071535.1 (AF255728) percent identity 100 in 451 aa"	/note="similar to GB:AA071535.1 (AF255728) percent identity 100 in 451 aa"	
/codon_start=1	/codon_start=1	
/evidence="not experimental"	/evidence="not experimental"	
/transl_table=1	/transl_table=1	
/product="chromosomal initiator protein"	/product="chromosomal initiator protein"	
/protein_id="BAC63096.1"	/protein_id="BAC63096.1"	
/db_xref="GI:28810158"	/db_xref="GI:28810158"	
/translation="MTEDNQIFNNRVLLEAOSQKQATTFEPVHDARLKVXKHIAITLDDMKLEFWKMLKDVILTAGEEVYNAQISVYVEEDIMKIQNQRKINQKQKQALNSLPTVDSKSESENFIOGDENRMAVASIAVANTPCTYTNPLTEIWGPGIGKTHLMAIGNSVLENDENARIKYTEAFENFEVYIHRIDTDMELKEKFNLDLIDIO SLAKKTLSTGSEEFNFENALHNNNKQIVLTSRPTPHLDLRVTRRKMLGTVIITPDEFRVALITNKIOEVNPIPDPTIEVLAOFDSNVDELGALKDISLVANFOIDTTVDIAAIAIRKQDGPKNIVIPLEIIOAOQKGYGVTAIEIKRTQNIVLARQVAMFLAREMTDNLPRIGKEFGGRDSTVLAVNKKIMNIDODESLEIEITIKNKIK"	/translation="MTEDNQIFNNRVLLEAOSQKQATTFEPVHDARLKVXKHIAITLDDMKLEFWKMLKDVILTAGEEVYNAQISVYVEEDIMKIQNQRKINQKQKQALNSLPTVDSKSESENFIOGDENRMAVASIAVANTPCTYTNPLTEIWGPGIGKTHLMAIGNSVLENDENARIKYTEAFENFEVYIHRIDTDMELKEKFNLDLIDIO SLAKKTLSTGSEEFNFENALHNNNKQIVLTSRPTPHLDLRVTRRKMLGTVIITPDEFRVALITNKIOEVNPIPDPTIEVLAOFDSNVDELGALKDISLVANFOIDTTVDIAAIAIRKQDGPKNIVIPLEIIOAOQKGYGVTAIEIKRTQNIVLARQVAMFLAREMTDNLPRIGKEFGGRDSTVLAVNKKIMNIDODESLEIEITIKNKIK"	
1742..2878	1742..2878	
/gene="SpS0002"	/gene="SpS0002"	
1742..2878	1742..2878	
/gene="SpS0002"	/gene="SpS0002"	
/note="similar to GB:AA096838.1 (AE009953) percent identity 99 in 378 aa"	/note="similar to GB:AA096838.1 (AE009953) percent identity 99 in 378 aa"	
/codon_start=1	/codon_start=1	
/evidence="not experimental"	/evidence="not experimental"	
/transl_table=1	/transl_table=1	
/product="putative DNA polymerase III beta subunit"	/product="putative DNA polymerase III beta subunit"	
/protein_id="BAC63097.1"	/protein_id="BAC63097.1"	
/db_xref="GI:28810159"	/db_xref="GI:28810159"	
/translation="MIQFSINRLEFIHALNATKRAISTKNAIPILSSIKIEVPTGVTITGNSQISIENTIPVSNENAGLITSPGAILLEASPFINISSLPDININKEIEOHQVLTSGSEITLKGKDQDQYPRLOEVSFENPLIILKRLKLSIAETAFAPASOESRPILTVGHVYLSNMHDKRAVADSHRMQRRLTLDNTSADPPVYIPKSLRESVFTDD IETVEFFSPQILFSEHISFTYRLLEGNPDRDLMTFEFEVEVFNQSLRHARERAFLISNATONGTVKLEITONHISAHNSPEVKVNEDDIVSOGSDLTISFPYTLIESIKAIKSETVVIHFLSPVRPFTLPGDEESFIQILTPVRIN"	/translation="MIQFSINRLEFIHALNATKRAISTKNAIPILSSIKIEVPTGVTITGNSQISIENTIPVSNENAGLITSPGAILLEASPFINISSLPDININKEIEOHQVLTSGSEITLKGKDQDQYPRLOEVSFENPLIILKRLKLSIAETAFAPASOESRPILTVGHVYLSNMHDKRAVADSHRMQRRLTLDNTSADPPVYIPKSLRESVFTDD IETVEFFSPQILFSEHISFTYRLLEGNPDRDLMTFEFEVEVFNQSLRHARERAFLISNATONGTVKLEITONHISAHNSPEVKVNEDDIVSOGSDLTISFPYTLIESIKAIKSETVVIHFLSPVRPFTLPGDEESFIQILTPVRIN"	
2953..3150	2953..3150	
/gene="SpS0003"	/gene="SpS0003"	
2953..3150	2953..3150	
/gene="SpS0003"	/gene="SpS0003"	
/note="similar to GB:AAK33148.1 (AE006472) percent identity 100 in 65 aa"	/note="similar to GB:AAK33148.1 (AE006472) percent identity 100 in 65 aa"	
/codon_start=1	/codon_start=1	
/evidence="not experimental"	/evidence="not experimental"	
/transl_table=1	/transl_table=1	
/product="conserved hypothetical protein"	/product="conserved hypothetical protein"	
/protein_id="BAC63098.1"	/protein_id="BAC63098.1"	
/db_xref="GI:28810160"	/db_xref="GI:28810160"	
/translation="MYOIGSFVEMKRPKPHACVIVKETGKKANQWVLRVGADIKIQCTNC OHVIMSRVDFERRIKKVLQP"	/translation="MYOIGSFVEMKRPKPHACVIVKETGKKANQWVLRVGADIKIQCTNC OHVIMSRVDFERRIKKVLQP"	
3480..4455	3480..4455	
/gene="SpS0004"	/gene="SpS0004"	
3480..4455	3480..4455	
/gene="SpS0004"	/gene="SpS0004"	
/note="similar to GB:AAK33149.1 (AE006472) percent identity 98 in 371 aa"	/note="similar to GB:AAK33149.1 (AE006472) percent identity 98 in 371 aa"	

/codon_start=1
 /evidence=not_experimental
 /transl_table=11
 /product="putative GTP-binding protein"
 /protein_id="BAC63109.1"
 /db_xref="GI:28810161"
 /translation="MALTVIGILPVPNGSKSTLEPNATTKGAAGAAVNPFIATIDPVGAVN
 EYVDRLOKATLEPLIPKPKVPTPTFEETDITAGLYKGSRGEGLNKFLANIREVDAIYH
 VYRAFDENYVKEGGREDAPVPVADITDINLELILADLESINKRTARVKMAKQDK
 KESVAEPVNLQIKVPLEEGKSARVILETDEEKKVYKGLLTKPVLVYANDEKVK
 ANPDIDIVYQJRDPAATENAAEVVISAAREBILEDDEDEKEFLAIGLITSGYDK
 LTRAYHLGICTYTPAGKEVYRAVTFKGIKAPQAAGITIHSDPEHGFIRAVTMSYDD
 LMTVSEKAIVKAGRLREBGRKVVYDGDGIMERRFNV"
 /length=5234
 /gene="SPS0005"
 /db_xref="GeneID:5234
 /note="similar to GB:AAK3150.1 (AE006472) percent
 identity 99 in 189 aa"
 /codon_start=1
 /evidence=not_experimental
 /transl_table=11
 /product="putative peptidyl-L-cRNA hydrolase"
 /protein_id="BAC63101.1"
 /db_xref="GI:28810162"
 /translation="WVKMTVGAGNPGSKTEKTKKHNIYGAINDIVKNKNDLPTDDKNF
 KAQISTINHEKVEKPTPTMNSGIVAKLLIYTNIDITDILIYDLDMEVSKL
 RLRSQSGVANGKGISIIAHITQOEFRNRKVGIRGLKGMTVNHMGORTEDNINAI
 SLTDVVYNAVYVLEQNDIEFKTQMGEFG"
 /length=8740
 /gene="SPS0006"
 /db_xref="GeneID:8740
 /note="similar to GB:AAK3151.1 (AE006472) percent
 identity 99 in 1167 aa"
 /codon_start=1
 /evidence=not_experimental
 /transl_table=11
 /product="putative transcription-repair coupling factor"
 /protein_id="BAC63101.1"
 /db_xref="GI:28810163"
 /translation="MDILIEFSQNKVKVQSHGSLTTLGRLVWGLSGSSKALAIASRY
 LDDOKIIVVSTQNGVIEKSLDDEELFQFFADVDAAEFTFMSMDALSTI
 ETLPLRNPSSGVNIEVLSIGARTLIPNDVFTSQIQLTVGADYSDTLTKQLMIG
 YOKVQSVLPGEFSFGDILDIYEITQELPYRLEFGDDIDSIRCFHPETQKSEFOLE
 GVFIIPASDILFEASDPOGILDIYEITQELPYRLEFGDDIDSIRCFHPETQKSEFOLE
 OSLAFKMSILDIYIPKGIPIFEDPOGILDIYEITQELPYRLEFGDDIDSIRCFHPETQK
 YQAFNRELRLHKRPATFSPNFKGILKIFDOMHQLTOYAMQEFNFOPPLIDELIKRY
 OKNOTTVIAQVESQTAPELSEKSFODYQRLPLVANSNQIVSRSSQIVIGAISSGFVA
 DETALITGEHETTKIRKRRASINSNAERKDYNELAVGDVYVNVHNGIGRFLGIE
 TITQIGHRDVTIYOQNSDRISLPIQDISLSKYVSADGKEPKINKLNDGRQKTKQ
 KVARQVEDIADLLLVGLAESRQSGFSFPDDLOKAFDDPFAFVFEEDQISLSEIKR
 ADMSGQMDRLDVLGVDGEGKEVGMRAAFKYNHKOVAVYPTVIAQOHENKRA
 RPNVYPEVDLSPRSRKEQAEETLERVAKGQIDITIGHRLSKIVYSDDLIIYD
 EEQRGAKHLEKLRKTRVDVLTLPATIPRLHMSMSGIRLVSIEPIVNRKVPQ
 TVYLENNAGKEALIREDRDGOGLPYVYNKVDTEKKVAEIQELPEASIGVGHQV
 SEIDIENTLIEFINDDYDLVATVITIEGVDSINWTFLETNADHMGSLTYLQRGV
 GRSNIEIAAYLAFIRNDKVLTESEKRELAIKGETELGSEFKIAMRLDSIGANILGA
 SOSGIRDSVGEPMYSLQLEOAIASROGKTIVROKGMTETNIOIDAVLPDIYADNRQK
 IDIYKRIEIOSREBYLNQDILLDRFGTVPQVAVLEIALIKRHMNDAFALVERK
 NNQVIAVREVSLSLTFLODFPEALSTHLKAKISEHQKIDIVFYVRHOKDRIILEE
 LMLFGERLSEIKIRKNSVFK"
 /length=9174
 /gene="SPS0007"
 /db_xref="GeneID:9174
 /note="similar to GB:AAK3152.1 (AE006472) percent
 identity 100 in 98 aa"
 /codon_start=1
 /evidence=not_experimental
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="BAC63102.1"

```

/db.xref="GI:28810164"
/translation="MKRIIVGEIMRLDKLYKVSRLIKRVSAREVADGKIRIVNGILAK
STNNKLKLDHEISFKNGLVTRVIEIKSTKKEKDLKMYELIISETRITLNEEA"
9161..9532
/gene="SPS0008"
9161..9532
/gene="SPS0008"
/note="similar to GB:AA196843.1 (AE009953) percent
identity 100 in 123 aa"
/codon_start=1
/translation="not experimental"
/translation="not experimental"
/product="putative cell division protein (DivIC)"
/protein_id="BAC63103.1"
/db.xref="GI:28810165"
/translation="MKKPSIVQIUNNHYIKKENILKKFEEESOKRNPFGIILVSMAF
LEILPTALVKSVDYDEPKQNOYVKKLKEYNELNSTKKEQGLNERLKDDFVAKKAR
AKYIISRGEMITVITPOLPK"
9529..9654
/gene="SPS0009"
9529..9654
/gene="SPS0009"
/note="unknown"
/codon_start=1
/evidence=not_experimental
/translation="not experimental"
/product="hypothetical protein"
/protein_id="BAC63104.1"
/db.xref="GI:28810166"
/translation="MMMDNIVLKITEAFASDKKLAEIQENQKVEESQYIQKN"

```

Alignment Scores:		
Pred. No.:	4.25e-108	311600
Score:	1862.00	379
Percent Similarity:	66.98%	Conservative: 124
Best local Similarity:	50.47%	Mismatches: 20
Query Match:	47.20%	Indels: 30
DB:	1	Gaps: 13

OY		20	SerLysAsnSerLysArg-----PheThrValThrLeuValGlyValPheLeuMetIle	37
Dd		111775	AGCGCTAACACCAACAGCAGCAAAACGACGATCCGATTACTCGAAGAATTTTTCACGTTT	111833
OY		38	PheAlaLeuValThrSerMetValGlValAlaLysThrValPheGlyLeuValGluSerSer	57
Dd		111835	GTAAGCTCATATGGAGATAGTAGTGTTCTTCATCACAGCGCTCGGAGCGTGAAGAAACAATCA	111894
OY		58	ThrProAsnAlaIleasnProAspSerSerSergIuTyTrArgTPPyGlyTYrgLusSer	77
Dd		111895	GTCGCCAAAT-----AAACAAGCTCAGTTCAGATTTTCGGTGTATGCCTATGATCT	111948
OY		78	TyrValArgGlyHisProTyrTyrlLysGlnPheArgValAlaHisAspLeuArgValAsn	97
Dd		111949	TATCTAAAGGTACCACACACTATAGTCGCTTAAAGACTTACCATTAATTTTAAAGTAAAT	112008
OY		98	LeuGluGlySerArgSerSeryrGlnValTyrcySPheAsnLeuLysValAlaPheProLeu	117
Dd		112009	TTAGATGGAAGTAAGAGATCAACACATACGTCTATTATTACAAAACACTTTCATCC	112068
OY		118	GlySerAspSerSerValLysLysTrpTyrlLysLysHisAspGlyIleSerThrLysPhe	137
Dd		112069	AAGTCGAAATGATTAGATCACAAAGATGATAAAAACTGAAGGACGTAAATGAACCTTT	112128
OY		138	GluAspTyrAlaMetSerProArqIleThrGlyAspGluLeuAsnGlnLysLeuArgAla	157
Dd		112129	ATCAAGATTACGATATAAACCAAGGATTACAGCGACAGCTTACAACAAATATATTGAGG	112188
OY		158	ValMetTyrAsnGlyHisProGlnAsnAlaAsnGlyIleMetGluGlyLeuGluProLeu	177
Dd		112189	ATTCTCTATAAAGGATATCCTAATGATATGTAATGAGGATAAAGAAAGGATATGATCTCTA	112248

OY	178	AsnAlaIleahgValThhGInGluAlaValTrrPyrTyrSerAspAsnAlaProIleSer	197
OY	178	AsnAlaIleahgValThhGInGluAlaValTrrPyrTyrSerAspAsnAlaProIleSer	197
Db	112249	AACGGTATTATTAGTACTCAAAATCTATTTGGTACTATACGATAGTGGTCTATATTTCT	112308
OY	198	AsnProAspGluSerPheLysArgGluSerGluSerAsnLeuValSerThrSerGlnLeu	217
Db	112309	GATACTACTAAAGCTTTTCACACAAAGAAACAGATCTTAAATTCATTCACACGACACTA	112368
OY	218	SerLeuMetArgGlnAlaLeuLysGlnLeuIleAspProAsnLeuAlaThrLysMetPro	237
Db	112369	CAACTAAATGCCGTAATGCCGTAATAAAGATTATATTCCAAAGAGAGGTACAAAGCTTACCG	112428
OY	238	LysGlnValProAspAspPheGlnLeuSerIlePheGluSerGluAspLysGlnLys	257
Db	112429	AATCGATGTTCCCGCAAAATTCAGTTAGTATTTTTCATTCCTTCGAT	112476
OY	258	TyrAsnLysGlyTyrGlnAsnLeuLeuSerGlyLysValProThrLysProProThr	277
Db	112477	-----AAGACTTTCCAAATCTTTTAAGTCTGCTAGTATGTTCCGGATCTCCCAAAA	112530
OY	278	ProGlyAspProPheMetProProAsnGlnProGlnThrThrSerValLeuIleArgLys	297
Db	112531	CCAGGAGAGAAG-----CTCTCGGCTAAAACTGAAAAAACATCACTGCTATTTCAGAAA	112584
OY	298	TyrAlaIleGlyAspTyrSerLysLeuLeuGlnGlyAlaThrLeuGlnLeuThrGlyAsp	317
Db	112585	TATCGCGAGGAGACTACTCTTAACTTCAGAGGAGCAACTTAAAGCTTGCTCAAAAT	112644
OY	318	AsnValAsnSerPheGlnAlaArgValPheSerSerAsnAspIleGlyLysValIleGlu	337
Db	112645	GAAGGAATGCTTTTCAGAGAAAAATTTTGTAGTAAACAGTCTGAGAAAGAAAGTAGAG	112704
OY	338	LeuSerAspGlyThrTyrThrLeuThrGlnLeuAsnSerProAlaGlyTyrSerIleAla	357
Db	112705	CTACTATATGGACTTACTGTTTAACTAGTACGCTAAACCTCCCAAGATATGAGTTGCC	112764
OY	358	GluProIleThrPheLysValGlnAlaGlyLysValTyrThrIleIle-----Asp	374
Db	112765	ACACCCTACTTCTTTAAAGTTCCACCTCGAGAAAGTT-----TTATATAAAATAAAGAA	112818
OY	375	GlyLysGlnIleGluAsnProAsnLysGlnIleValGluProTyrSerValGlnAlaTyr	394
Db	112819	GGTCATTGTTGAGAACACCAAAATAGGAATCGACAGACCCTATTTCTGTAACGCTTTC	112878
OY	395	AsnAspPheGluGlnPheSerValLeuThrThr---GlnAspTyrAlaLysPheTyrTyr	413
Db	112879	AATGACTTGAAGAGATATGGTATTATTACTGATTTTAATACATAGTGAAGTTTACTAT	112938
OY	414	AlaLysAsnLysAsnGlySerSerGlnValValTyrCysPheAsnAlaAspLeuLysSer	433
Db	112939	GCAAAAAATACTAAAGCAACTATACAGTTGCTACTGTTTCAATCTGATTTTACACTCA	112998
OY	434	ProProAspSerGluAspGlyLysThrMetThrProAspPheThr---ThrGlyGlu	452
Db	112999	CCACTGACTATATGATATCAGCGACCAATATATGATCCGATGTCAGTAAAGGTAAGAG	113058
OY	453	ValLysTyrThrHisIleAlaGlyArgAspLeuPheLysTyrThrValLysProArgAsp	472
Db	113059	ATAAAGTATACACATGTTCTTGCGTATGATTTGTATAATATGACCCACACCAACAGAC	113118
OY	473	ThrAspProAspThrPheLeuLysHisIleLysLysValIleGluLysGlyTyrArgGlu	492
Db	113119	AAAGATGCTGATTTCTCTTAAACATATCAAGAAATCTTGATTAAGGTATTAAAGAG	113178
OY	493	LysGlyGlnAlaIleGlyTyrSerGlyLeuThrGlnThrGlnLeuArgAlaAlaThrGln	512
Db	113179	AAGGGGATACC-----TATTAAGACATTTAACTGAAAGCTCAAGTTTAGACACCAACCAA	113232
OY	513	LeuAlaIleTyrTyrPheThrAspSerAlaGluLeuAspLysAspLysLeu-----	529
Db	113233	TTAGCTATTATTACTATATACAGATATGCTGTATTAAACAACACTTAAGCATATATATGAT	113292
OY	530	---LysAspTyrHisGlyPheGlyLysPheTaaAspSerThrLeuAlaValAlaLysIle	548

DB	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
Db	113293		AE009963	11420 bp	DNA	linear	BCT 03-APR-2002
QY	549	LeuValIGluTyrAlaGlnAspSerAsnProProGlnLeuThrAspLeuSphPheIle	AE009963	11420 bp	DNA	linear	BCT 03-APR-2002
Db	113353	TTGATCACCCTTATGCTGAGGATGTCACCTTACCANTAGCTCAAAATCTAGATTCTTTGTT	AE009963	11420 bp	DNA	linear	BCT 03-APR-2002
QY	569	ProAsnAsnAsnLysTyrGlnSerLeuIleGlyThrGlnThrPheIleProGluAspLeuVal	AE009963	11420 bp	DNA	linear	BCT 03-APR-2002
Db	113413	CCTAATAGCAGATAGATATACGAGCATTATTTGGACGACGATGATCATTAATGATGATT	AE009963	11420 bp	DNA	linear	BCT 03-APR-2002
QY	589	AspIleIleArgMetGluAspLysLys---GluValIleProValIleThrHisAsnLeuThr	AE009963	11420 bp	DNA	linear	BCT 03-APR-2002
Db	113473	GATGTTATTCTTATGAGAAATTAACAGCGCTCCATTATTCCTTATCTCAAGATTAACT	AE009963	11420 bp	DNA	linear	BCT 03-APR-2002
QY	608	LeuArgLysThrValThrGlnLeuAlaGlyAspArgThrLysAspPheHisPheGluIle	AE009963	11420 bp	DNA	linear	BCT 03-APR-2002
Db	113533	ATTCTTAAACGTTTACTGGAACCTATTGCAGATAGCAAAAAAGAAATTTAACTTTGAATA	AE009963	11420 bp	DNA	linear	BCT 03-APR-2002
QY	628	GluLeuLysAsnAsnLysGlnGlnLeuLeuSerGlnThrValLysThrAspLysThrAsn	AE009963	11420 bp	DNA	linear	BCT 03-APR-2002
Db	113593	CATTAAATCTCTCGATGAGCACAAGCTAATAGGAAACATTCGCAAACTCTGGAGAA	AE009963	11420 bp	DNA	linear	BCT 03-APR-2002
QY	648	LeuGluPheLysAspGlyLysAlaThrIleAsnLeuLysHisGlyLysLeuThrLeu	AE009963	11420 bp	DNA	linear	BCT 03-APR-2002
Db	113653	CTCACAGTTACAGATGAGAAACCTACCTCACATTAAGGATGAGAAATCATTTAGTTT	AE009963	11420 bp	DNA	linear	BCT 03-APR-2002
QY	668	GlnGlyLeuProGlnGlyTyrSerTyrLeuValGlyGlnThrAspSerGluGlyTyrLys	AE009963	11420 bp	DNA	linear	BCT 03-APR-2002
Db	113713	GAGGGGCTTACCTTACGGTTCCTTATGAATTTACGAAAGGGGCTTCCAGATTATGAG	AE009963	11420 bp	DNA	linear	BCT 03-APR-2002
QY	688	ValLysValAsnSerGlnGluValAlaAsnAlaThrValSerLysThrGlyIleThrSer	AE009963	11420 bp	DNA	linear	BCT 03-APR-2002
Db	113773	CTAAGTGTTAATGGAATAATGCACAGATGGAAAGCAGAAACCTCAGTTAAGAA	AE009963	11420 bp	DNA	linear	BCT 03-APR-2002
QY	708	AspGluThrLeuAlaPheGlnAsnAsnLysGluProValIleProThrGlyLysAlaSerGln	AE009963	11420 bp	DNA	linear	BCT 03-APR-2002
Db	113833	GATGAGACTGTGCTTTTGAAACCGAAAGATCTTGTCACCACTGCTTT-GACAAC	AE009963	11420 bp	DNA	linear	BCT 03-APR-2002
QY	728	LysIleAsnGlyTyrLeuAlaIleValIleAlaGlyIleSerLeuGlyIleThr---746	AE009963	11420 bp	DNA	linear	BCT 03-APR-2002
Db	113892	AGATGG---GGCTATCTTATCTTGTGTTGTTTACTTG-TCCTATTGGGTTATGTTT	AE009963	11420 bp	DNA	linear	BCT 03-APR-2002
QY	747	GlyIleHisThrIleArgIleArgLysHisAsp757	AE009963	11420 bp	DNA	linear	BCT 03-APR-2002
Db	113948	GGCTATTGGTCGTAAAGGCACTAAATAATGCAC113980	AE009963	11420 bp	DNA	linear	BCT 03-APR-2002
RESULT 7							
LOCUS							
DEFINITION							
ACCESSION							
VERSION							
KEYWORDS							
SOURCE							
ORGANISM							
REFERENCE							
AUTHORS							
TITLE							
JOURNAL							
MEDLINE							
PUBMED							
REFERENCE							
AUTHORS							

Porcella, S.F., Parkins, L.D., Beres, S.B., Campbell, D.S., Smith, T.M., Zhang, Q., Kapur, V., Daly, J.A., Yeasty, L., George, J., and Musser, J.M.
 Direct Submission
 Submitted (31-JAN-2002) Laboratory of Human Bacterial Pathogenesis/Rocky Mountain Laboratories/NIH/NIH, 903 S. 4th St., Hamilton, MT 59840, USA
 Location/Qualifiers
 1. 11420
 /organism="Streptococcus pyogenes MGAS8232"
 /mol_type="genomic DNA"
 /strain="MGAS8232"
 /db_xref="taxon:186103"
 107. .502
 /gene="ssb"
 /note="synonym: spym18_0120"
 107. .502
 /gene="ssb"
 /note="best blastp match gb|AAK04490.1|AE006276_5 (AE006276) single-strand binding protein [lactococcus lactis subsp. lactis]"
 /product="putative single-strand binding protein"
 /protein_id="AAJ96931.1"
 /db_xref="GI:19747395"
 /translation="MYNKVIAIGRLVAKPELVKATDMDKHVARTSLAVNRFFKNASGR EAGFISVYVVGKLAETLVSYAKSLSDIGELRTKRYDKDGVHYTEVLQCSFOLL ESRARMRNRYNDLVLDLEEDLPF"
 complement(757. .1398)
 /gene="spym18_0121"
 complement(757. .1398)
 /note="best blastp match gb|AAC97156.1| (U49397) unknown [Streptococcus pyogenes]"
 /codon_start=1
 /transl_table=11
 /product="hypothetical protein"
 /protein_id="AAJ96932.1"
 /db_xref="GI:19747396"
 /translation="MLIVLAGTIGAGKSSLAALGHEHIGDVFVEAVDNNPVLDIYIO DPKKVAFLQIYFLNKRPKSIKRAYQADNNILDRSIFEDFLKLNKKNVKTIELD IYQELANNLEELGMPKKRPDLITIDVSFDMLEIKERKRSFQVDNPSLEQIY HOVHGEPWYEDYKVSPPKMKIDGNSLDFVQNPQDLATVLMKIDTKLKLHL"
 complement(1418. .2395)
 /gene="spym18_0122"
 complement(1418. .2395)
 /note="best blastp match gb|AAC97155.1| (U49397) unknown [Streptococcus pyogenes]"
 /codon_start=1
 /transl_table=11
 /product="hypothetical protein"
 /protein_id="AAJ96933.1"
 /db_xref="GI:19747397"
 /translation="WTKLNSFRIGDVEIPIHRTVLAPMACVTSAPRTTAKFAGVIV MMSIEKGLLYNNKTLHMLHIDENHEPMSIQLFQDAGLKRADFIOITNKRAJIV DIMGCPVNVVYVNEAGAKLRDPDKITIHVKEVSLDPLTKMRGTGADSSLAWE NALAESAGVYALAMGRTRQEQYGTGCHDETARVSKATIKIPIFGNGPVSVOAK FMEIEITGAVMTGRAMNPPYLFQIINHEFTGQELPDLPFAKKIDIAKDHKLRLIN LKGETIYVREFRILAPHYLRGTGAAKAVKRAVSRATLEAVEAIFETVR"
 complement(2382. .3254)
 /gene="spym18_0123"
 complement(2382. .3254)
 /note="best blastp match sp|Q9ZB45|HSLO_STRPY_33_KDA CHAPERONIN (HEAT SHOCK PROTEIN 33 HOMOLOG) (HSP33)"
 /codon_start=1
 /transl_table=11
 /product="putative HSP33"
 /protein_id="AAJ96934.1"
 /db_xref="GI:19747398"
 /translation="MDKIITKSIAGSCAFRAYVUDSTETVALAQEKINTLSSSTVALGR TLIANQILANOKGDSKITVYKVIIGDSSFGHIIISVADTKGHVKGVIQNTGVDIKTATG

EVLVGPFMGNGHFVITIDYGTGNPYSTPLTNGEIGEDFAVYLTSEEQTPSAIGLV LLEBNKRVKAVGFQVOLFEGASEEELARERKLOMPAISHLAKNHYDALLEAIY GDEPKRISEEPJLFOODCSREREFEALMLLPKADQAMIDEDKAEIVCOFCGTXYO FNSDLEALINDKA"
 complement(3380. .4000)
 /note="nra.2 (spym18_0124): sequence confirmed upstream mutation gene fragment; best blastp match gb|AAC97147.1| (U49397) Nra [Streptococcus pyogenes]"
 complement(4262. .4915)
 /gene="nra.1"
 /note="synonym: spym18_0125"
 complement(4262. .4915)
 /gene="nra.1"
 /note="best blastp match gb|AAC97147.1| (U49397) Nra [Streptococcus pyogenes]"
 /codon_start=1
 /transl_table=11
 /product="N-terminal fragment of Nra"
 /protein_id="AAJ96935.1"
 /db_xref="GI:19747399"
 /translation="MPYVKKKSDFLVETRYLEQSIKDSKSELVLLFKSPITIFSHVAK QRGTLTAQVLYKKEKLDPPFGNNLDITIKKRTICCFVVPKREYLAHQYDSTIKL LVFINKGTSQPLIKFSKRYFLSSSARLRSLLIKLREGLRSKNTYGEETRI RYLIALMYSKFGIYIPLDLMQIIRFLSQSATNLRTSPWLEEFSEYNNMLALS"
 5345. .6919
 /gene="cbp"
 /note="synonym: spym18_0126"
 5345. .6919
 /gene="cbp"
 /note="best blastp match gb|AAC97148.1| (U49397) Cpa [Streptococcus pyogenes]"
 /codon_start=1
 /transl_table=11
 /product="putative collagen binding protein"
 /protein_id="AAJ96936.1"
 /db_xref="GI:19747400"
 /translation="MQKRDKTNGSANNRKQTTIGLVFLTFVALIGIVGFSIRAF GAEQSTETKRTSVIIRKYAEGYSKLLBQATLKQIISGSGQESFSSSGOKIQ LSGDTIIFETKSPQGEIAPETPTFTAGKVTIKGDOGFVNOKEVAEPYSVAY NDDPSGFINPKPFPYKGYKAYKANNQSOVYVCNVLDLHSPDSLDGGETDPFN EKEETIYTHLGLDLSYANNPASINDELISQYKVVLEKGRVDSITANLTSVEFR AATQVLIYFTDSVDLNDLADVHGFEALTEALNAKRETVAAEDSANLPNSLDFV VPNSKYSQSLIGIQVHPRESLVDIIRMEDQAPITPTHKLTISKVTYGTIADKKEFN FEIHLKSSDGOAISGTPTNSGELVTDKAPFTLLDGSLLVEGLPSGSYEITFTG ASDYEVSNGKNAPDGAATKASVKEDETTTEFRKQVLVPTGTLTGAIYIMLLLV LIGLWVLIGRKGKAKND"
 6912. .7433
 /gene="lepa"
 /note="synonym: spym18_0127"
 6912. .7433
 /gene="lepa"
 /note="best blastp match gb|AAC97149.1| (U49397) Lepa [Streptococcus pyogenes]"
 /codon_start=1
 /transl_table=11
 /product="lepa"
 /protein_id="AAJ96937.1"
 /db_xref="GI:19747401"
 /translation="MTNINLRLENELFKAFIRLVAKISITIGFGLTLPQYIFGWMI LNVKSPALASAGGILYRITDRIYHINDVYVENDTLVGRIVAOAGGEVSFQDGG LINGHPPEKEVGYLYPHSSGPNFPYKVPETGYFTLNDYREBRDLSRYGALPINOI KGRISTYLIRVRGI"
 7454. .8488
 /gene="spym18_0128"
 7454. .8488
 /gene="spym18_0128"
 /note="best blastp match gb|AAC97152.1| (U49397) unknown [Streptococcus pyogenes]"
 /codon_start=1
 /transl_table=11
 /product="hypothetical protein"
 /protein_id="AAJ96938.1"
 /db_xref="GI:19747402"

ORGANISM Streptococcus pyogenes
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Streptococcus.

REFERENCE 1 (bases 1 to 2845)
AUTHORS Sela,S., Avily,A., Tovli,A., Burstein,I., Caparon,M.G. and Hanski,E.
TITLE Protein F: an adhesin of Streptococcus pyogenes binds fibronectin
via two distinct domains
JOURNAL Mol. Microbiol. 10 (5), 1049-1055 (1993)
MEDLINE 95020565
PUBMED 7934855

COMMENT Original source text: Streptococcus pyogenes (strain JMS75) DNA.
FEATURES
source Location/Qualifiers
1..2845
/organism="Streptococcus pyogenes"
/mol_type="genomic DNA"
/strain="JMS75"
/db_xref="taxon:1314"
425..2404
/gene="prtF"
/gene="prtF"
/function="binds fibronectin"
/codon_start=1
/transl_table=1
/product="protein F"
/protein_id="AA26964.1"
/db_xref="GI:425480"
/translation="MEVRSGEKMNKKIEFLNKEASFLAHTKKRRBPVAVTLGVFPMILAC
AGALGGVAVAYADEKTVPSHSSNPPEFPGYDAYCKEYGNIMTYRDLAVNLNG
SRSYOYCFNICSNIYPSQKNSFLKNMKFKLEGKSKSVDAHTTKLKELEBORLSL
LYNATPMDANGIKGGLHNAITVOTVAHNYSDNSOYQETTLSEKAKGKISRSQV
LTHREALKLIDNLEATAVNKIPSGRLNIFESNAYONLSAEVVPDDPKPGST
SEHNPTPELDGPIPEDPKHPDNLPEPLPYALDGEVPEVSESLPELPKLE
LDQEVPEKSIDLPLEPRYEFNNKQDPSLAGSEETVITEVYGQONPVIDKLE
PNTEGSGNVEFEDTRKEPVLNMGQSESEVETKPDQTMGSGOTTPOVETDKPEPV
LNGQSESEVETKPDQTMGSGOTTPOVETDKPEPVLMGQSESEVETKPDQTMGSG
OTTPOVETDKPEPVLMGQSESEVETKPDQTMGSGOTTPOVETDKPEPVLMGQ
SESEVETKPDQTMGSGSEVETATVDFPVDNEPRVEENRKRPNKITPLIP
ATGDIENVALFELILSLVLSFLKNKSNKKV"
1327..1457
/gene="prtF"
/note="repeat domain 1"
1631..2170
/gene="prtF"
/note="repeat domain 2"
BASE COUNT 977 a 472 c 616 g 780 t
ORIGIN

Alignment Scores:
Pred. No.: 1.33e-31 Length: 2845
Score: 633.50 Matches: 203
Percent Similarity: 42.62% Conservative: 129
Best Local Similarity: 26.06% Mismatches: 236
Query Match: 16.06% Indels: 211
DB: 1 Gaps: 30

US-09-494-297-2 (1-757) x STRPRTNF (1-2845)

QY 20 SerLysAsnSerLysArgPheThrValThrLeuValGlyValPheLeuMetIlePheAla 39
Db 494 ACAAAAAGAAAAGCGCATTTGCTGCACCTTTAGTGGAGATCTTTTATTAGCTTTTGCGCA 533

QY 40 LeuValThrSerMetValGlyAlaLysThrValPheGlyLeuValGluSerSerThrPro 59
Db 554 TCGCGGGGTGCTATTCGGTTTGGTCAAGTAGCCTATCGTCGCGATGAGAACTGTCGCCG 613

QY 60 AsnAlaIleAsnProAspSerSerSerGlyTyrArgTyrTyrGlyTyrGluSerTyrVal 79
Db 614 AGTCATAGTAGTCCA-----AATCCGAGCTTTCCTGCTACGCGTTATGATGCTTATGCT 667

QY 80 ArgGlyHisProTyrTyrLysGlnPheArgValAlaHisAspLeuArgValAsnLeuGlu 99
Db 668 AAGCAATATCTCGATATATATTTGGACTGTTATCATGACTTTCAGAGTAATTTAAAT 727

QY 100 GlySerArgSerTyrGlnValTyrCysPheAsnLeuLysValAlaPheProLeuGlySer 119
Db 728 GCGAGTCGAGCTACCAAGTTTATGTTTAAATATTAATCTATTTATCTACCAAAAA 767

QY 120 AspSerSerValLysLysTyrTyrLysHisAspGlyLysSerThrLysPheGluAsp 139
Db 788 AATAGTTTCATAAAAAATGGTTTAAAAAATGAGGAATGAGCAATCTTGTGATGAT 847

QY 140 TyrAlaMetSerProAlaGlyIleThrGlyAspGlyLeuAsnGlnLysLeuAlaValMet 159
Db 848 TATGCTCACACTACTAACTAAGTAAAGAAAGAGTGGAGAGGCTCTTAAGTTACTG 907

QY 160 TyrAsnGlyHisProGlnAsnAlaAsnGlyIleMetGlnLysLeuGluProLeuAsnAla 179
Db 908 TATATGCGTATCCCAACGATGCTAATAGTTATATGAAGACTGGAGCTTAAATGCT 967

QY 180 IleArgValThrGlnGluAlaValTyrPyrTyrSerAspAsnAlaProIleSerAsnPro 199
Db 968 ATTACTGTACACAATATGAGATTGGCATTTACTACATATTCACAATATCAATTT 1024

QY 200 AspGluSerPheLysArgGlyLysGluSerGluSerAsnLeuValSerThrSerGlnLeu 219
Db 1025 GAGACATTATGGAGTCTGAGCGAAGAAAGATTAAGTTAGTGCACAACTTACTTTA 1084

QY 220 MetArgGlnAlaLeuLysGlnLeuLysAspProAsnLeuAlaThrLysMetProLysGln 239
Db 1085 ATGCGTAGGACATTGAAAAACATATGATGCCAATTTAAACCTACTGACGTTAATTA 1144

QY 240 ValProAspAspPheGlnLeuSerIlePheGluSerGluAspLysGlyAspLysTyrAsn 259
Db 1145 ATCCCAACAGCCTATCTTTAAATATCTTTGAGCTCGAA-----AAT 1186

QY 260 LysGlyLysThrGlnAsnLeuSerGlyLysLeuValProThrLysProProThrProGly 279
Db 1187 GAAGCTTACCAAAATCTTTTAAGTCTGAATATGATGATGATGATGATGATGATGATG 1246

QY 280 AspProPheMetProProAsnGlnProGlnThrThrSerValLeuIleArgLysTyrAla 299
Db 1247 GAG-----ACGTCACAGACATATTCCTAAACCTCCGAG----- 1279

QY 300 IleGlyAspTyrSerLysLeuLeuGluGlyAlaThrLeu-----GlnLeuThr 315
Db 1280 -----TTGGATGCGCATCTCAATTCACGAGCAACCAACATCA 1318

QY 316 GlyAspAsnValAsnSerPheGlnAlaArgValPheSerSerAsnAspIleGlyGluArg 335
Db 1319 GATGATATTTAAGCTACATGCGCGCGATGCTAGAT-----GGTGAAGA 1369

QY 336 IleGluLeuSerAspGlyThrTyrThrLeuThrGluLeuAsnSerProAlaGlyTyrSer 355
Db 1370 GTC-----CCAGAAAGTTCACAGCGACGACTTAGAGCTT 1402

QY 356 IleAlaGluProIleThrPheLysValGluAlaGlyLysValTyrThrIleLeuAspGly 375
Db 1403 GCGCTTCCCTCCCATTAATGCCAGAG-----TTAGATGGT 1435

QY 376 LysGlnIle---GluAsnProAsnLysGluIleValGluProTyrSerValGluAlaTyr 394
Db 1436 CAAGAAAGTCCCTAAACCTACTATGACTTA-----CCTTTGAAGTCTCTGCTTAT 1489

QY 395 AsnAspPheGluGluPheSerValLeuThrThrGlnAsnTyrAlaLysPheTyrTyrAla 414
Db 1490 GAG----- 1492

QY 415 LysAsnLysAsnGlySerSerGlnValValTyrCysPheAsnAlaAspLeuLysSerPro 434
Db 1493 -----TTTAACATATAAGCCATACCT 1516

QY 435 ProAspSerGluAspGlyGlyLysThrMetThrProAspPheThrThrGlyGluValLys 454
Db 1517 CTAGCGGGGTAG-----TGTGTGAGACGGAG 1543

OY	455	Tyr---ThrlsIleAlaGlyArgAspLeuPheLysTrpValLysProArgAspThr	473
Db	1544	TATATTACCGAAGCTGTATGGAAATCAA-----CAGAACCCGTGGTAAATT	1588
OY	474	AspProAspThrPheLeuLysHisIleLysLysValIleGluLysGlyTyrArgGluLys	493
Db	1589	GATGAAA-----AAACTCCGAATGGAACAGGTTTT-----TCA	1621
OY	494	GlyIleAlaIleGlyTyrSerGlyLeuThrGluThrGluLeuArgAlaIleThrGluLeu	513
Db	1622	GGAAATATGTTGGAGACAGAAAGATACGAAAGCAGACAGCTGTTGATGGAGGCCAA---	1678
OY	514	AlaIleTyrTrpPheThrAspSerAlaGluLeuAspLysAspLysLysAspTyrHis	533
Db	1679	-----AGTGAAGCTCTGTTGAATTTACTTAAGACACTCAACAGCATGAGT	1723
OY	534	Gly-----PheGly	536
Db	1724	GGCTCAAACTCCTCAGCTTGACACAGACATACGAAAGCCAGAACTGTTGATGGGA	1783
OY	537	AspMetAsnAspSerThrLeuAlaValAlaLysIleLeuValGlyTyrAlaGlnAspSer	556
Db	1784	GTGCAAGTAGTCT-----GTTGAATTTACTTAAGATTAAT	1819
OY	557	-----AsnProGluLeuThrAspLeuAspPheIlePro	569
Db	1820	CAACAGCATGATGCTGTCACAAACCTCCAGATTGAGACAGAAACACGAAAGAGCA	1879
OY	570	AsnAsnAsnLysTyrGlnSerLeuIleGlyThrGlnTrpHisProGluAspLeuValAsp	589
Db	1880	GAA-----	1882
OY	590	IleIleArgMetGluAspLysGluValIleProValThrHisAsnLeuThrLeuArg	609
Db	1883	---GTGTTGATGGAGAGCCAAAGTAGCTGTGAATTTACT-----AAA	1924
OY	610	LysThrValThrGlyLeuAlaGlyAspArgThrLysAspPheHisPheGluIleGluLeu	629
Db	1925	GACACTCAACAGCATGATGAGTGTGTCACAAACACTCCT-----CAGGTTGAGACA	1972
OY	630	LysAsnAsnLysGln---GluLeuLeu-----SerGlnThrValLysThrAsp	644
Db	1973	GAAATATACGAAAGAGCCAGAGATGTTGATGGAGAGCTCAAACTGAGTCTGTGAATTTACT	2032
OY	645	LysThrAsnLeuGluPheLysAspGlyLysAlaThrIleAsnLeuLysHisGlyGluLeu	664
Db	2033	AAAGATACTCAACAGCATGATGAGTGTGTCACAAACTCCTCAGATTGACACGAAAGACGC	2092
OY	665	LeuThrLeuGlnGlyLeuProGluGlyTyrSerTyrLeuValLysGluThrAspSerGlu	684
Db	2093	AAAGAGCCAGAAAGTGTGATGGAGGTCGCAAGTGATGCTGTGAATTTACT-----	2143
OY	685	GlyTyrLysValLysValAsnSerGlnGluValAlaAsnAlaThrValSerLysThrGly	704
Db	2144	-----AAAGACCTCAACAGCATGAGTGTGTCACAGGAAACGCG	2185
OY	705	IleThrSerAspGluThr-----LeuAlaPheGluAsnAsnLys-----	717
Db	2186	ACTGTTGTTGAGATACCGCTCCCAAGTTAGTGTTCATTTTGCAATATGAGACCCANA	2245
OY	718	-----GluProValAlaPro---ThrGly	724
Db	2246	GTGGAAGACGATCGGGAAGAAACCTACAAATAATTAACACCTATTCCTTCGCAAGAGA	2305
OY	725	ValAspGlnLysIleAsnGlyTyrLeuAlaLeuIleValIleAlaGlyIleSerLeu	743
Db	2306	GATATTGAGAAATGTTTGGCCCTTCTTGGAATCCTTATTGTCAGTACTTTCATTT	2362
RESULT 9			
AF009908	2801 bp	DNA	linear
LOCUS			BCT 13-FEB-1998
DEFINITION			Streptococcus pyogenes fibronectin/fibrinogen binding protein F (pf15) gene, complete cds

ACCESSION	AF009908
VERSION	AF009908.1
KEYWORDS	GI:2271466
SOURCE	'
ORGANISM	Streptococcus pyogenes Streptococcus pyogenes Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Streptococcus.
REFERENCE	1 (bases 1 to 2801)
AUTHORS	Katerov,V., Andreev,A., Schalen,C. and Totolian,A.A. Protein F, a fibronectin-binding protein of Streptococcus pyogenes, also binds human fibrinogen: isolation of the protein and mapping of the binding region
JOURNAL	Microbiology 144 (Pt 1), 119-126 (1998)
MEDLINE	98129085
PUBMED	9467904
REFERENCE	2 (bases 1 to 2801)
AUTHORS	Katerov,V.E., Andreev,A.S., Schalen,C. and Totolian,A.A. Direct Submission
TITLE	Submitted (20-JUN-1997) Med. microbiol., Lund University, Solvegatan 23, Lund 22362, Sweden
JOURNAL	Location/Qualifiers
FEATURES	1..2801
source	/organism="Streptococcus pyogenes" /mol_type="genomic DNA" /strain="EPI949" /serotype="M15" /db_xref="taxon:1314" 450..2507 /gene="prtF15" 450..2507 /gene="prtF15" /codon_start=1 /transl_table=11 /product="fibronectin/fibrinogen binding protein F" /protein_id="AAC38155.1" /db_xref="GI:2271467" /translation="MPLRGEKMNKKIFLKEASFLATTKRRRRPAVTLVGFFMLFAC AGSLGFGVAAADEKTEPNRINSNPEFPWGYSGRLIARYHLNVNSKEQVAC ENLRPEPKRESSSPNNMKYKLDSGETFEPPKYAENPRPSEGLRHRIILVYNGPS MEIKGDILALITVNATIMVYSADAPINDINPEFTSEANDINIPPOOLTMEPEAR KLISDENLVQVPNSNRKLSIFESSDKSYONLLSAEYVPDDPPRPGRGTSEHNPRTPEL DGTIPEDPKRPDESSGPALPPLMPLELDGEVEPVPSLEPALPVLPLEDGEVEVE VPSSLPEPALPPLPELDGEVEPVPSLEPALPPLPELDGEBEIPEVPSLEPAL PLPMLELDGEVEPEKPSIDLPLEVRYEFNNKDSPLAGESGETDYII EYVGNQNFV TDIKKLPEPTGFSGSNVETERTKEPGVLMGGOSSEVFETKTDTQAGMSGOTTPOVETEED TKRGVLTMGGSSESVETKTQTOTMSQTAPOVETEDTKREPVLMMGGOSSEVFETKT QAGMSGOTAPOVETEDTKREPVLMMGGOSSEVFETKTDTQTSSTGSETVSIVIEDTRPL VFHDNDNEPKVEEHRERKPYKNITPILPATGDLENVAFLDLILSLVSLSKNKON NKV"
gene	
CDS	
BASE COUNT	954 a 489 c 597 g 761 t
ORIGIN	
Alignment Scores:	
Pred. No.:	3.62e-30 Length: 2801
Score:	611.00 Matches: 210
Percent Similarity:	41.70% Conservative: 124
Best Local Similarity:	26.22% Mismatches: 239
Query Match:	15.49% Indels: 228
DB:	Gaps: 33
US-09:494-297-2 (1-757) x AF009908 (1-2801)	
OY	20 SerLYSAsenSerLYSArCPheThValrThrLeuValAlglyAlpheUmEllePhaA 39
Db	519 ACAAAAGAAAAGGCCGATTTCCTTGCGTCACTTAGTGGCAGCTTTTAAAACTTTTCCA 578
OY	40 LeuVALTrSerMetValAlglyAlaLyThrValrPheglYLeuValAlguSerSerThrPro 59
Db	579 TGC GGCGGTCTATCG GTTTGGTGCACGATGACCTATGCGTGGCATGAGAAGACTGACCG 638
OY	60 AsnAlaIleAsnPTrOAspSerSerSergLUtYArgrPrpyrGLyTYrgUSeTytyVal 79
Db	639 AATCGTATAATAT-----TCAAATCCGACGTTTCCCTTGGAATVGGTATGATTCGATAAGC 692

```

QY      80  ATGGLYHISPROTYRTRYLYSGINPHEARGVALALANISASPLEUARGVALASINLEUGLU
Db      693  GGAAGATGTGAAGATAT-----CACATCTGAAGAAAT----- 728
QY      100  GLYSERASERTYRGINVALTYRCYSRHEASINLEULYSVALAPHEPROLEUGLISER
Db      729  GGAAGTAGAGAGATAGCGTATGTTTAACTGAAAGCTTTAGCGCAAAAAGAA 788
QY      120  ASPSERVALYSYSTRPTRYLYSLYSNHSASGLYTLSESRTHRYSRPHGLUAP
Db      789  GAAAGTAGACAGCTTATGTTGATTAATAATAGATGGAGTACAGAAACTTTTAAAAA 848
QY      140  TYRALMETSERPROARGILETHRGLYASPGULLEASNGLINLSLEUARGALVALMET
Db      849  TAGCGTGAAGAACCTGAGTTTACTGAGAGGAATTAAGACAGACATATATGAAGGTTCTC 908
QY      160  TYRASNGLYHISPROGLINASINLAASNGLYILEMETGLUGLYLEUGLUPROLEUASINLA
Db      909  TACAAATGATATCTCT--AATAGTAAATGAATATAGAGGGATAGACCTCTCAACGCT 965
QY      180  ILEARGVALTHRCINGLUALAVALTRPTRYTRYSERASPSANLAPROILESERASINPRO
Db      966  ATCTTAGTACACAAATGCGCATATGCTATTACTCGGATRGTCCAGCATTAATGATAT 1025
QY      200  ASPGLUSERPHELYSARGLUSERGLUSERASINLEUVALSERTHRISERGLINLEUSERLEU
Db      1026  AATAACTCTCTTACTAGTAGAGCCATGATCTTAATATACCACTCAGACGATACGTTA 1085
QY      220  METARGLALALEULYSGLINLEULE-----ASPPROASINLEUALATHIRYSMETPRO
Db      1086  ATGCGTGAAGCTCTGTGAAAACTAATTTCTCTGTGAAAAATTTGGTT----- 1133
QY      238  LYSGLINVALPROASPSRPHENISERGLINLEUSERILEPHEGLUSERGLUSPLYSGLYASPLYS
Db      1134  AAACAAGTTCACACGACTTAACCTTAGATTTTGAAGCTCTGAT----- 1181
QY      258  TYRASINLSGLYTRYGINASINLEULEUSERGLYGLYLEUVALPROTHRYSPROPROTHR
Db      1182  -----AGCTCCACCAAAATCTTTAAGTGTGATATGATGATGATGATGATGATGAT 1235
QY      278  PROGLYASP----- 280
Db      1236  CCTGGTATACGTCAGACATATACTCTAAACTCCGAGTGTGATGCGACATCCAAATCCC 1295
QY      281  -----ProPROMETPROPROASINLEUINPROGLIN 289
Db      1296  GAGGACCCAAACGTCACAGATGAGACTTCAGACCTCGCGCTTCCCATTTGATGCCAGAG 1355
QY      290  THRTHRSERVALLEULEARGLYSTRALALEGLYASPTYSERLYSLEULEUGLUGLY
Db      1356  CTAGATGTGAAGAGAGTCCAGAAAGTTCCAAAGCGAGACTTAGAACCTGGCGTTCGCCCA 1415
QY      310  ALATHLEUGLINLEUTHRGLYASPSANVALASINSERPHEGLINLAIRYVALPHESESER
Db      1416  TTGGTCCCAAGCTCATATGCTGAAGAAGTCCCAAGAACTTCCAAAGC----- 1460
QY      330  ASNAPDILEGLYGLUARGILEGLULEUSER-----ASPGLY 341
Db      1461  -----GAGAGCTTAGAAGACTCGCGCTTCCCATTAATGCCAGACGTAGATGGC 1508
QY      342  THRTRYTHLEUTHRGLEULASINSESRPROALAGLYTRYSERILEALAGLUPROILETHR
Db      1509  ---CAAGAGTCCCAAGACTTCCAAAGCGAGAGCTTAGAACCTCGGCTCTCTCATTAAG 1565
QY      362  PHELYSVALGLUALAGLYSVALTYRTHIRLE----- 372
Db      1566  CCAGAGCTAGACGCTGAAAGAAATCCCAAGAGGAGACTTAGAACCTCGCGCTT 1625
QY      373  -----ILEASPGLYLSGLINLE-----GLUASINPROASINLSGLINLE 385
Db      1626  CCTCCATTATGCCAGAGTATGATGTCAAGAAAGTCCCTGAATAAAACCTAGATTTGACTTA 1685

```

```

QY      386  VALGLUPROTRYSERVALGLUALATYRASNASPHEGLUGLUPHESESERVALLEUTHRTHR
Db      1686  ---CCTATGCAAGTCTCTCGTATGAG----- 1709
QY      406  GLINASITRYRALALYSPHERITYRTRALALYASINLSYASNGLYSERSEGLINVALTYR
Db      1709  ----- 1709
QY      426  CYSRHEASINLAASPLEULYSSESRPROASPSERGLUASPOLYGLYLSYTHIRMETTHR
Db      1710  ---TTTAACAAATAAACAACCAATCACTTAGCGGGTAG----- 1745
QY      446  PROASPSRPHETHRTHRGLYGLUVALTYR---THRHSILEALAGLYARG----- 461
Db      1746  -----TCTGTGAGACGAGATTTATATATGCAAGTCAATGGAATCAACAGAAC 1793
QY      462  ---ASPLEUPHELYSTRYTHIRVALLYSPROARGASPTHRASPROKASPTHRPHELEU
Db      1794  CCGTGTGATATGATTAATAAAACTT-----CCGAATGAAGACGTTTTCAGCAAAATAG 1847
QY      480  LYSNHSILEULYSVALILEGLULYSGLYTRYARGLULYSGLYGIN-----ALALILE 497
Db      1848  GTTGACACAGAAAGACGAAAGAGCCAGAGGTGATGGAGGCCAAAGTAGTGTCTGT 1907
QY      498  GLUTYRSERGLYLEUTHRGULUTHRGINLEUARGALALATHRCINLEUALATHRYTYR
Db      1908  GAATTTACGAAGAACACTCAACGACGACGACGCGCTCAACA----- 1949
QY      518  PHETHRASPSERLALALEULASPLYSASPLYSLEULYASPRYTHISGLYPHEGLYASP
Db      1950  ---ACTTCCTCAGGTATGACACAGAGACGAAAGACCCAGAGCTGTG---ATGGAGCT 2003
QY      538  METASNASPSERTHRLEUALAVALALALYSLILELEUVALGLUTYRZALINASPSERASN
Db      2004  CAAAGTAGAGCT-----GTGAATTTACGAAGACACT--- 2036
QY      558  PROPROGLINLEUTHRASPLEUASPPHEPHEILEPROASINASINLSYTYRGLINSERL
Db      2037  -----CAACAGGCGC 2045
QY      578  ILEGLYTHRGINTRPHISPROGLINASPLEUVALASPLILEARGMETGLUSPLYS
Db      2046  ATAGTGTCAACACACCTCTCAG-----GTTGACACAGACAGACAGAA 2090
QY      598  GLUVALILEPROVALTHRHISASINLEUTHRLEUARGLYSTRVALTHRCYLEUALAGLY
Db      2091  GAG-----CCAGAAATG-----TTGATGGGA 2111
QY      618  ASPARGTHRILYSASPSRPHENISRHEGLUIILEGLULEULYSASINLSYGLINLEULEU
Db      2112  GGCCAAGTAGCTCTGTGAATTTAG-----AAAGACACTCAACGACGAGCATGAGC 2162
QY      638  SERGLINTHRVALLYSTRHRSPLYSRHSINLEUGLUPHELYASPGLYLSYALATHIRLE
Db      2163  GGTCAACAGCT-----CCTCAGGTTGAACAGACAGACGAAAGCCAGAA 2210
QY      658  ASNLEULNHSIGLYLUSERLEUTHRLEUGLNGLYLEUPROGLINLYTRYSERTRYLEU
Db      2211  GTCCTTAGAGAGGTCAAGTAGCTGTGATTAATTAACAAGAT----- 2255
QY      678  VALLYSGLUTHRASPSERGLUGLYTRYTRYVALYLSVALASINSESRGLINLUVALALASN
Db      2256  ---ACTCAAAAGAGCTCGAGCGGTTTCAGTGAAGACGTAGACATTTGATATAGCGCT 2312
QY      698  ALATHIRVALSERLYSTRHGLYIETHRSERASPSGLUTHRLEUALAHEGLUASINLSY
Db      2313  CCGAAGTTA-----GTGTCCATTTTGAACATAATAGAG 2345
QY      718  -----GLUPROVALVALPRO----- 722
Db      2346  CCCAAAGTGGAAGAGATGGGAAAGCCTACAAAAAATATACACTTATCTCTCTGCA 2405
QY      723  THRGLYVALASPOLINLSYLSINLSNGLYTRYLEUALALEUILEVALILEALAGLYLESER 742

```

Db 2406 ACAGGAGATATGTAGATGTTTGGCCTTCTCTTGGAATCCATTATTTGTGCTACTTTCT 2465
Oy 743 Leu 743
Db 2466 ATT 2468

RESULT 10
SPSFB 2263 bp DNA linear BCT 13-JUL-1995
LOCUS S.pyogenes sfb gene for fibronectin-binding protein (partial).
DEFINITION X67947.S42389
ACCESSION X67947.1 GI:511149
VERSION fibronectin-binding protein; sfb gene.
KEYWORDS Streptococcus pyogenes
SOURCE Streptococcus pyogenes
ORGANISM Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Streptococcus.

REFERENCE 1 (bases 1 to 2263)
AUTHORS Talay, S.R.
TITLE Direct Submission
JOURNAL Submitted (31-JUL-1992) S. R. Talay, GBF, Mascheroder Weg 1, 3300
BRAUNSCHWEIG, FRG
REMARK 2 (bases 1 to 1030)
AUTHORS Talay, S.R., Valentini-Weigand, P., Jerlstrom, P.G., Timmis, K.N. and
TITLE Chhatwal, G.S.
Fibronectin-binding protein of Streptococcus pyogenes: sequence of
the binding domain involved in adherence of streptococci to
epithelial cells
JOURNAL Infect. Immun. 60 (9), 3837-3844 (1992)
MEDLINE 92363585
PUBMED 1386839
REFERENCE 3 (bases 1 to 2263)
AUTHORS Talay, S.R., Valentini-Weigand, P., Timmis, K.N. and Chhatwal, G.S.
TITLE Domain structure and conserved epitopes of sfb protein, the
fibronectin-binding adhesin of Streptococcus pyogenes
JOURNAL Mol. Microbiol. 13 (3), 531-539 (1994)
MEDLINE 95089690
PUBMED 7527894
REFERENCE 4 (bases 1 to 2263)
AUTHORS Talay, S.R.
TITLE Direct Submission
JOURNAL Submitted (13-JUL-1994) S. R. Talay, GBF, Mascheroder Weg 1, 3300
BRAUNSCHWEIG, FRG
COMMENT On Jul 17, 1994 this sequence version replaced gi:47433.
FEATURES
source Location/Qualifiers
1..2263
/organism="Streptococcus pyogenes"
/mol_type="genomic DNA"
/strain="DSM 2071"
/db_xref="taxon:1314"
124..2040
/gene="sfb"
124..2040
/gene="sfb"
/codon_start=1
/transl_table=1
/product="fibronectin-binding protein"
/protein_id="CAA48133.1"
/db_xref="GI:511150"
/translation="MNKMKFLNKESAGLATKKRRRAVLGVFPMILSAGLIGRG
QVAVADEKTVPRVSONPEFPMGIDFTVGPYTRHNLQNLNGSKTYOYCPNLK
FEPKKEGSYFPMNKKRWDSEETFKVYADPRKDNESRYIDVELKNIKRLVINGP
NNGINGIEGLPLNALIVTONAVYSDNSIFNTDNFTTEADLNKIKPEGLIMAV
ALKKLIDPKLSESLKVPSTFLNIFESODKLYONLSAEVNEPDKGEPPEHGP
KTPLEDGTPIDEGORNESLEPTLPVMDGQVPEVPSLEPALPMLPELDGQV
VPEVPSLEPALPMLPELDGQVPEVPSLEPALPMLPELDGQVPEVPSLEPALP
YITEVGNQNPVDIKLPLNETGSGNMVETEDTKEPGLVLMGQSEVETKQTQVG
MSQTPPOVETEDTKEPGLVLMGQSEVETKQTQVMSQTSQVETEDTKEPGLVLM
GQSEVETEDTKEPGLVLMGQSEVETKQTQVMSQTSQVETEDTKEPGLVLMGQ
SEVETEDTKEPGLVLMGQSEVETKQTQVMSQTSQVETEDTKEPGLVLMGQSG
FSTVIVEDTRPKLVFHDNNEPKVEENREKPKNTITPILPATGDIENVLAFGLIIS

sig_peptide 124..270
BASE COUNT 790 a 383 c 477 g 613 t
ORIGIN

Alignment Scores:
Pred. No.: 9.3e-30 Length: 2263
Score: 603.00 Matches: 208
Percent Similarity: 44.17% Conservative: 133
Best Local Similarity: 26.94% Mismatches: 235
Query Match: 15.29% Indels: 197
DB: 1 Gaps: 34

US-09-494-297-2 (1-757) x SPSFB (1-2263)

Oy 20 SerLysAnsSerLysArgPheThValThrLeuValGlyValPheLeuMetIlePheAla 39
Db 172 ACANAAAGAAAAAGCCGATTTGCTGCTACCTTACTGGAGCTTTTATCTCTTGGCA 231
Oy 40 LeuValThrSerMetValGlyAlaLysThrValPheGlyLeuValGluSerSerThrPro 59
Db 232 TCCCGGGGTGCTATCGGTTTGGTCAAGTACCTATGCTGCCGATGAGAAAGCTGCT 291
Oy 60 AsnAlaIleAsnProAspSerSerSerGluThrArgThrProGlyLysGluSerThrVal 79
Db 292 CATAGAGTTACT-----CAAAATCCTGAGTTTCCCTGATGTTGTTGATTTCTAC-- 342
Oy 80 ArgGlyHisProLysThrLysGlnPheArgValAlaHisAspLeuArgValAsnLeuGlu 99
Db 343 AAAGGA---CCTTACACAGATAT-----CATATTACAGCTAAATCTCTAC 387
Oy 100 GlySerArgSerLysGlnValLysCysPheAsnLeuLysLysAlaPheProLeuGlySer 119
Db 388 GGAAGTAAACATATTCAGGCTTATGTTTATTTAAAGATTTGAACCTTAAGAAAGAA 447
Oy 120 AsperSerValLysLysTrpLysLysLysHisAspGlyLysSerThrLysPheGluAsp 139
Db 448 GGTGTTATTTTCCGAATGTTGATTAAGAGGAGGAGTACTGAGACACATTTGTAAA 507
Oy 140 TyrAlaMetSerPro-----ArgIleThrGlyAspGluLeuAsn 152
Db 508 TATGCTGATTAACCCAGAAAGATTAAGAAAGTCCAGAGTTAAGATGTAATTTGAA 567
Oy 153 GlnLysLeuArgAlaValMetLysAsnGlyHisProGlnAsnAlaAsnGlyIleMetGlu 172
Db 568 AAAATATATTAAGAGGCTATATTAAGATGATGCAATAACGGTAAATGGAATCATGAG 627
Oy 173 GlyLeuGluProLeuAsnAlaIleArgValThrGlnGluAlaValIleThrLysSerAsp 192
Db 628 GGTTAGACCATTAATGCTATCTTGTACCTAAGAAAGTGGTGGTACTATTCGAT 687
Oy 193 AsnAlaProIleSerAsnProAspGluSerPheLysArgGluSerGluSerAsnLeuVal 212
Db 688 AATTCATCATATTTTAAATACGATACGATCTTTACGATGAGTAAAGATCTTAATAATA 747
Oy 213 SerThrSerGlnLeuSerLeuMetArgGlnAlaLeuLysGlnLeuIleAspProAsnLeu 232
Db 748 AATCCTGAGCAATTAATTAATGCTGCTGCTTAAAGAAATGATGATGATCAAGTTA 807
Oy 233 AlaThrLysMetProLysGlnValProAspAspPheGlnLeuSerIlePheGluSerGlu 252
Db 808 TCTAAGACATATTAATTAATGCTGCTGCTTAAAGAAATGATGATGATCAAGTTA 867
Oy 253 AspLysGlyAspLysLysLysGlyLysGlnAsnLeuLeuSerGlyLysLeuValPro 272
Db 868 GAT-----AACCTATATCAAAATGCTTAAAGGCTGAGTGTGTACCA 909
Oy 273 ThrLysProProThrProGlyAspProProMetProProAsnGlnProGlnThrSer 292
Db 910 GAAATCCCTTAAACCTGCTGAG-----ACCCCTGACATGAGCCCAAAATCCCGAG 963
Oy 293 ValLeuIleArgLysTyrAlaIleGlyLysPyrSerLysLeuLeuGluGlyAlaThrLeu 312

```

Db      964 TTG-----GATGCACCTCAATCTCCGAAGCCCA-----993
QY      313 GlnLeuThrGlyAspAsnValAsnSerPheGlnAlaArgValPheSerSerAsnAspIle 332
Db      994 CAACGTCGAATAGTTTAAAGACCTACGCTCCACCACTG-----1035
QY      333 GlyLysArgIleGluLeuSerAspGlyThrTyrThrLeuThrGluLeuAsnSerProAla 352
Db      1036 -----ATGCTAGACGGA---CAAGAGTCCCGAGAGTCCCAAGTGAAGC 1077
QY      353 GlyTyrSerIleAlaGluProIleThrPheLysValGluAlaGlyValTyrThrIle 372
Db      1078 TTGAACCTGCGCTCCCTCCATGTATGACAGCTAGATGCGCAAGAGTCCCAAGATT 1137
QY      373 -----1LeaspGlyLysGln 377
Db      1138 CCAGGAGAGCTTAGAACCTGCGCTCCCATTTGATGCGCAGATGATGATGTCAGAA 1197
QY      378 Ile---GluAsnProAsnLysGluIle---ValGluProTyrSerValGluAlaTyrAsr 395
Db      1198 GTCCCGAATAAACCTGATGTTGACTTACCTATGAAAGATCCTGTTATGAGTTAACAAT 1257
QY      396 AspPheGluGlnPheSerValLeuThrThrGlnAsnTyrAlaLysPheTyrAlaLys 415
Db      1258 -----AAAGACCAATCACCCTTAGCGGGTGAAGTGTGAGAGAGATATTATTCGAA 1311
QY      416 -----AsnLysAsnGly 419
Db      1312 GTTATGGAATCAACAGAACCTGTGATATTGATAAAACCTCCGAATGAACAGCT 1371
QY      420 SerSerGlnValTyrCysPheAsnAlaAspLeuSerProProAspSerGluAsp 439
Db      1372 -----TTTTCAGAAATATGTTTACAGACAGAGATAGCAAAAG 1410
QY      440 -----GlyGlyLysThrMetThrProAspPheThr-----Thr 450
Db      1411 CCAGAGATGTTGATGGAGGCTCAAGTGAAGTCTGTTGAATTTACTAAAGACTCAACA 1470
QY      451 Gly-----GluValLysTyrThrHisIleAlaGly---460
Db      1471 GGCATGATGTCGTCAAACAACTCCTCAGAGTTGAGACAGAACACAGAAAGCCAGAGGTG 1530
QY      461 -----ArgAspLeuPheLysTyrThrValLysProArgAspThrAspPro 475
Db      1531 CTGATGGAGGTCGAAGTGAAGTCTGTTGAATTAC-----AAAGTACTCAACA 1581
QY      476 AspThrPheLeuLysHisIleLysValIleGluLysGlyTyrArgGluLysGly---494
Db      1582 GGCATGATGTCGTCAAACAGCTTCTCAGGTTGAGACAGAACATCAGAAAGCCAGAGGTG 1641
QY      495 -----GlnAlaIleGluTyrSerGlyLeuThrGluThrGlnLeuArg 508
Db      1642 CTGATGGAGGCGCAAGAGTCTGTTGAATTTACTAAAGACACTCAACAGGCAATGAGT 1701
QY      509 AlaAlaThrGln-----LeuAlaIleTyr 516
Db      1702 GGTCAACAACAACCTCCTCAGTTGAGACAGAAAGACAGCAAGAGCAGAGTATGATGCGA 1761
QY      517 TyrPheThrAspSerAlaGluLeuAspLysAspLysLeuLysAspTyrHisGlyPheGly 536
Db      1762 GGTCAAGATGAATCTGTTGAATTTACTAAAGACACTCAACAGGCAATGAGCGGTTTCAGT 1821
QY      537 AspMetAsnAspSerThrLeuAlaValAla-----LysIleLeuValGlu 551
Db      1822 GAA-----ACAGTGACCAATGTTGAAGATACCCGTCGGAAGTTAGTTCCAT 1869
QY      552 TyrAlaGluAspSerAsnProGlnLeuThrAsp-----563
Db      1870 TTT-----GACATATATAGCCCAAGTCGAGAGAGATCGGAAAAAGCTTACAAAAAAT 1923
QY      564 LeuAspPhePheIleProAsnAsnAsnLysTyrGlnSerLeuIleGlyThrGlnTyrPhe 583

```

```

Db      1924 ATAAACCTATCTCTCTGC--AACAGAGATATTGAGATGTTTGGCCTTCTTGGAAT 1982
QY      584 ProGluAspLeuValAspIleIleArgMetGluAspLysGlyGluValIleProValThr 603
Db      1983 CCT-----TATTTTGTCAGTACTTCC 2003
QY      604 HisAsnLeuThrLeuArgLysThrValThrGlyLeuAlaGlyAspArgThrLysAspPhe 623
Db      2004 TATTTTATGCTTTTAAACAAACAAACAAACATTAAGTCGATCGT---AAAAGCTT 2060
QY      624 HisPheGluIleGluLeuLysAsnAsnLysGlnGluLeuLeuSerGlnThrValLysThr 643
Db      2061 CCATAGCAAAATA--ATGCCCAATATGCGAATAGTAGTAACCTGCTACGTCAGCAACA 2117
QY      644 AspLysThrAsnLeuGluPheLysAspGlyLysAlaThrIleAsnLeuLysHisGlyLys 663
Db      2118 -----AGTITTAATTCANAA-----2132
QY      664 SerLeuThrLeuGlnGlyLeuProGluGlyTyrSerTyrIleuValLysGluThrAspSer 683
Db      2133 -----GGTACTACTTAT-----ACATCT 2150
QY      684 GluGlyTyrLysValLysValAsnSerGlnGluVal 695
Db      2151 GATAATAATATAGCTGACAGATATTCAGACCAAGTA 2186

```

```

RESULT 11
GCU31115
LOCUS      2066 bp      DNA      linear      BCT 27-AUG-1996
DEFINITION      Group G streptococcus fibronectin binding protein (gfba) gene,
complete cds.
VERSION      U31115
KEYWORDS      U31115.1 GI:950168
SOURCE      Streptococcus sp. 'group G'
ORGANISM      Streptococcus sp. 'group G'
REFERENCE      1 (bases 1 to 2066)
AUTHORS      Kline,J.B., Xu,S., Bisno,A.L. and Collins,C.M.
TITLE      Identification of a fibronectin-binding protein (gfba) in
pathogenic group G streptococci
JOURNAL      Infect. Immun. 64 (6), 2122-2129 (1996)
MEDLINE      96239026
PUBMED      8675316
REFERENCE      2 (bases 1 to 2066)
AUTHORS      Kline,J.B., Xu,S., Bisno,A.L. and Collins,C.M.
TITLE      Direct Submission
JOURNAL      Submitted (06-JUL-1995) James B. Kline, Microbiology & Immunology,
University of Miami, 1600 NW 10th Ave, Miami, FL 33136, USA
FEATURES
source
1..2066
/organism="Streptococcus sp. 'group G'"
/mol_type="genomic DNA"
/strain="1750"
/db_xref="taxon:1320"
252..1994
/gene="gfba"
252..1994
/gene="gfba"
/note="gfba"
/codon_start=1
/transl_table=11
/product="group G streptococcal fibronectin binding
protein"
/protein_id="AB06623.1"
/db_xref="GI:950169"
/translation="MFLRGEKMMKMKFLNKEAGFWHTKKRRPAVTLGVFPMLLACA
GATGEGVAVYADEKTPVNPSPDPYPAWGDAYGAFLRHDLNVNLEGSPYQVY
CRLVYQDSKYNKGFPRKREKYDGNAYAEKQAAAPRYIDGDIENITLVNYGKPS
DANGIKGIDRTNAILVTOIATWYSKNSKFLVDEQKSELKRGELISEQVKLMRAL
RKLIISPDEKTYKNQLPENYKLNLFISKNSNIONLSTYVODNLOKPEDEKESPK
EENPKIPDLDGHEIPVPEPSPVPIPVILDGEVPEVSPSESLPELPLPMEIDGQ

```


EVPEKPSIDLPVEVRYEFNNKQDPSLAGESETEYITEVIGNOONPVDIDKLPNET
GFSGNMVEDTKEPEVLMGSGSESSEVETKQIOTGMSGOTTPOVEETKEGVLMG
QSESEVEFKROTMSGOTPOVEETKEGVLMGSGSEVEFKROTMSGSE
ATVEEDRPKLVFHFNDNNEPKVEENQEPKNTPILEPATDIENVLAFGLIILSVL
PIFSLKNKNNKKV

BASE COUNT 724 a 326 c 435 g 581 t
ORIGIN

Alignment Scores:

Score: 1.13e-24 Length: 2066
Percent Similarity: 523.00 Matches: 170
Best Local Similarity: 44.21% Conservative: 97
Query Match: 28.15% Mismatches: 233
DB: 13.26% Indels: 105
Gaps: 18

US-09-494-297-2 (1-757) x GGU31115 (1-2066)

QY 20 SerIysAsnSerIysArgPheThrValThrLeuValGlyValPheLeuMetIlePheAla 39
Db ACACAAAGAAAAGGCGATTGCTGCTACCTTAGGGAGAGCTTTTATGCTTTGGCA 377
QY 40 LeuValThrSerMetValGlyAlaLysThrValPheGlyLeuValGlySerSerThrPro 59
Db TGCGGGGCTGCTACGCTTTGGTCAAGTACCTATGCTGCCGATGAGAAGACTGTGCCG 437
QY 60 AsnAlaIleAsnProAspSerSerSerGluThrArgTrpArgTrpGlyGlySerIysVal 79
Db AATTTTAAAGCCCGAT-----CCAGATTATCCCTGTATGCTGATGATGCTTATCT 491
QY 80 ArgGlyHisProTrpTrpGlyGlnPheArgValAlaHisAspLeuArgValAsnLeuGlu 99
Db GGAGCGCTTTTAAGTAT-----CATGATTTAAGTTATTTGGAA 533
QY 100 GlySerArgSerTrpGlnValTrpCysPheAsnLeuLysAlaPheProLeuGlySer 119
Db GTTGTACAGCCCTTACCAAGTTATTTGCTTAAATTAGTAAGACAGACCGGTAAGGTT 593
QY 120 AspSerSerValLysTrpTrpTrpTrpTrpTrpTrpTrpTrpTrpTrpTrpTrpTrp 139
Db AATGCTTTACAGAAATTCGGGTTCAAAAAGTTGATGATTAAGTGTGTTTAAACAA 653
QY 140 TyrAlaMetSerProArgIleThrGlyAspGluLeuAsnGlnLysLeuArgAlaValMet 159
Db TATGAGCTATTCCTAGAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 713
QY 160 TyrAsnGlyHisProGlnAsnAlaAsnGlyIleMetGluGlyLeuGluProLeuAsnAla 179
Db TACATGATGATTCCAAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 773
QY 180 IleArgValThrGlnGluAlaValTrpTrpTrpTrpTrpTrpTrpTrpTrpTrpTrp 199
Db ATTTTATGATCTCAACTGCGATTTGGTATTTCTGATTAATCTCAACTTATATA---GTG 830
QY 200 AspGluSerPheLysArgGluSerGluSerAsnLeuValSerThrSerGlnLeuSerLeu 219
Db GATGACAGCTGAGTGAATTAAGATGATGATGATGATGATGATGATGATGATGATGAT 890
QY 220 MetArgGlnAlaLeuLysGlnLeuLysProAsnLeuAlaThrLysMetCProLysGln 239
Db ATGGCTGATGCTCTAGAAAATTAATATACACAGATTTTAAAGAAAACGTCACAAAATAC 950
QY 240 ValProAspPheGlnLeuSerIlePheGluSerGluAspGlyLysAspLysTrpAsn 259
Db TTGGCTGAGAAATTAATGTTAAATCTTTTATTTTCAAAAGAT----- 992
QY 260 LysGlyTrpGlnAsnLeuLeuSer----- 267
Db AACTGATTTCAAAATTTATTTAGTACTAGTATGATGATGATGATGATGATGATGATGAT 1052
QY 268 -----Gly 268
Db GAAGAACCTAAAGAGAGTCCAAAAGAGAACCCCAAAATTCAGATTTAGATGGG 1112

QY 269 GlyLeuValProThrLysProProThrProGlyAspProPheProAsnGlnPro 288
Db 1113 CATGAATCCCGAAGACTTCCACAGAACCTTCAGACCCAGTTATCCACCG----- 1163
QY 289 GlnThrThrSerValLeuIleArgLysTyrAlaIleGlyAspTrp----SerLysLeu 307
Db 1164 -----GTTATTTTGGATGCTGACAGAAAGTCCAGAAAGTTCACAGCAGCTTA 1211
QY 308 GluGlyAlaThrLeuGlnLeuThrGlyAspAsnValAsnSerPheGlnAlaArgValPhe 327
Db 1212 GAACCTGCGCTTCCCTCACTTATATGCCAGAG-----TTA 1244
QY 328 SerSerAsnAspIleGlyGluArg-----IleGluLeuSerAspGly 341
Db 1245 GATGCTCAAGAGTCCCTGAAAACCTGATGATGATGATGATGATGATGATGATGATGAT 1304
QY 342 ThrTrpThrLeuThrGluLeuAsnSerProAlaGlyTrpSerIleAlaGluProIleThr 361
Db 1305 GAGTTTAAACATTAAGACCAATCCGCTGAGGGGTGATGCTGACAGAGATATTT 1364
QY 362 PheLysValGluAlaGlyLysValTrpThrIleIleAspGlyLysGlnIleGluAsnPro 381
Db 1365 ACCGAGCTATGGAATCAACAGAACCCCTGTGATGATGATGATGATGATGATGATGAT 1418
QY 382 AsnLysGluIleValGluProTrpSerValGluAlaTrpAsnAspPheGluGluPheSer 401
Db 1419 AATGAACACAGCTTTTTCAGGAATATGGTGTGACAGACAGATGACGAAGAGCCAGACTG 1478
QY 402 ValLeuThrThrGlnAsnTrpAlaLysPheTrpTrpAlaLysAsnLysAsnGly---Ser 420
Db 1479 TTGATGGAGAGCGCAAGTGAAGTGTGATGATGATGATGATGATGATGATGATGATGAT 1538
QY 421 SerGlnValValTrpCysPheAsnAlaAspLeuLysSerProProAspSerGluAspGly 440
Db 1539 GGTCAAAACACTCTCATGCTGACAGACAGATGACGAAGAGCCAGAGCTGTGAGGA 1598
QY 441 GlyLysThrMetThrProAspPheThrThrGlyGluValLysTrpThrHisIleAlaGly 460
Db 1599 GGTCAAGTGAAGTCTGTTGAATTTACG-----AAAGATCTCAACAGGCTGATGATGT 1652
QY 461 ArgAspLeuPheLysTrpThrValLysProArgAspThrAspProAspThrPheLeuLys 480
Db 1653 CAA-----ACACCTCCAGGTGACACAGA 1679
QY 481 HisIleLysValIleGluLysGlyTrpArgGluLysGlnAlaIleGluTrpSer 500
Db 1680 GATACGAAA-----GAGCCAGAGAGTGTGATGAGGAGGCAAGT---GAGTCTGT 1727
QY 501 GlyLeuThrGluThrGlnLeuArgAlaIleThrGlnLeuAlaIleTrpTrpPheThrAsp 520
Db 1728 GGATTTACGAAGAC-----ACTCAACACAGCATATGAGTGTTCAGTGA 1772
QY 521 SerAlaGluLeuAspLysAspLysLeuLysAspTrpTrpHisGlyPheGlyAspMetAsnAsp 540
Db 1773 ACAGCGACTGTTGTTGAATATCGGCTCGCAAGTATGATGATGATGATGATGATGATGATGAT 1832
QY 541 SerThrLeuAlaValAlaLysIleLeuValGluTrpAlaGlnAspSerSerProGln 560
Db 1833 CCCAAA-----GTGGAAGGAATACGAGAAAACCTACAAAAT 1871
QY 561 LeuThrAspLeuAspPhePheIleProAsnAsnAsnLysTrpGlnSerLeuIleGlyThr 580
Db 1872 ATACACCTATC-----CTTCTGC-AACAGAGATATGGAATGTTTGGCCTT 1921
QY 581 GlnTrpHisPro 584
Db 1922 TCTTGAAATCCT 1933

RESULT 12
AF009914 AF009914 321 bp DNA linear BCT 13-FEB-1998
DEFINITION Streptococcus pyogenes fibronectin binding protein F gene, partial

ACCESSION	AF009914	cds.
VERSION	AF009914.1	GI:2267181
KEYWORDS		
SOURCE		
ORGANISM	Streptococcus pyogenes	
	Streptococcus pyogenes	
	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;	
REFERENCE		
AUTHORS	1 (bases 1 to 321)	
TITLE	Katerov,V., Andreev,A., Schalen,C. and Totolian,A.A.	
	Protein F, a fibronectin-binding protein of Streptococcus pyogenes,	
	also binds human fibrinogen: isolation of the protein and mapping	
	of the binding region	
JOURNAL	Microbiology 144 (Pt 1), 119-126 (1998)	
MEDLINE	98129085	
PUBMED	9467904	
REFERENCE	2 (bases 1 to 321)	
AUTHORS	Katerov,V.E., Andreev,A.S., Schalen,C. and Totolian,A.A.	
TITLE	Direct Submission	
JOURNAL	Submitted (23-JUN-1997) Medical Microbiology, Lund University,	
	Solvegatan 23, Lund 2262, Sweden	
FEATURES	Location/Qualifiers	
source	1..321	
	/organism="Streptococcus pyogenes"	
	/mol_type="genomic DNA"	
	/strain="M4 strain Umea"	
	/db_xref="taxon:1314"	
	<1..>321	
	/codon_start=2	
	/transl_table=1	
	/product="fibronectin binding protein F"	
	/protein_id="AAC38161.1"	
	/db_xref="GI:2267182"	
	/translation="FGQVATSADEKTYTPNKRSPDPYPMGIYAYGCGYEDISKTY	
	HLDEVNLNGSOVIYQVCFNIQKIFPYNVKSVYQKMFKAENSDFTPYAMPNVRQGE	
	ELSG"	
BASE COUNT	107 a 48 c 68 g 98 t	
ORIGIN		
Alignment Scores:		
Pred. NO.:	1.63e-09	Length: 321
Score:	272.50	Matches: 53
Percent Similarity:	64.81%	Conservative: 17
Best local Similarity:	49.07%	Mismatches: 31
Query Match:	6.91%	Indels: 7
DB:	1	Gaps: 2
US-09-494-297-2 (1-757) x AF009914 (1-321)		
OY	51 PheGlyLeuVal-----GlusSerSerThrProAsnAlaIleasnProAsp	65
	:: ::	
Db	2 TTGTGTCAGTAGCGCTATTCGCGGAGTGAAGAAGCTGCGCGATTTTAAAGCCGAT	61
OY	66 SerSerSerGluYrAArgTTPrYrGlyTylGluSerTyValAArgIyNHSProTyrTr	85
	::: :: :: ::	
Db	62 -----CCAGATTATTCCTGGTATGATTATGCTATGTTAGGTAAAGGATACCTCGATAT	115
OY	86 LysGlnhEarGValAlaIleHisAspLeuArgValasnuLeuGlySerAArgSerTyGln	105
	::: ::	
Db	116 GATATCAGTAGATTTATTCATGATTTTAAGGTTAAATCTAAACGGAAGTCAGGTCACAA	175
OY	106 ValTyrcysPheasnuLeuIySLySAlAphProLeuGlySerAspSerSerValIySlys	125
	:: ::	
Db	176 GTTATATGCTTATATATCAAAAATTTTCCGACAAATGTGAAGAACGTTACTACAAAA	235
OY	126 TrpTyrlsYlyshisAspGlyIleSerThrLysPheGluAspTyralaMetSerProArg	145
	::: ::	
Db	236 TGGTTTAAAAAAGCCGAGGGAATTCGATACCTTTTGACCTTATCCAAATGAATCCAGA	295
OY	146 IleThrGlyAspGluLeuAsnGln 153	
	::: ::	
Db	296 GTCCAAAGGAGAACTTAGTCAA 319	

RESULT 13	AF009913	313 bp	DNA	linear	BCT 13-FEB-1998
LOCUS	AF009913				
DEFINITION	Streptococcus pyogenes fibrinectin binding protein F gene, partial cds.				
ACCESSION	AF009913				
VERSION	AF009913.1	GI:2267179			
KEYWORDS					
SOURCE	Streptococcus pyogenes				
ORGANISM	Streptococcus pyogenes				
REFERENCE	Bacteria; Firmicutes; Lactobacillales; Streptococaceae; Streptococcus.				
AUTHORS	1. (bases 1 to 313)				
TITLE	Katerov,V., Andreev,A., Schalen,C. and Totolian,A.A.				
JOURNAL	Protein F, a fibrinectin-binding protein of Streptococcus pyogenes, also binds human fibrinogen: Isolation of the protein and mapping of the binding region				
MEDLINE	Microbiology 144 (Pt 1), 119-126 (1998)				
PUBMED	98129085				
REPERE	9467904				
AUTHORS	2 (bases 1 to 313)				
TITLE	Katerov,V.E., Andreev,A.S., Schalen,C. and Totolian,A.A.				
JOURNAL	Submitted				
FEATURES	Submitted (23-JUN-1997) Medical Microbiology, Lund University, Solvegatan 23, Lund 22362, Sweden				
Source	Location/Qualifiers				
CDS	1..313				
	/organism="Streptococcus pyogenes"				
	/mol_type="genomic DNA"				
	/strain="M4 strain 281C"				
	/db_xref="taxon:1314"				
	<1..>313				
	/codon_start=2				
	/transl_table=11				
	/product="fibrinectin binding protein F"				
	/protein_id="AAC38160.1"				
	/db_xref="GI:2267180"				
	/translation="FEGVAVASADEKTVPNFKSPDPYPMWGYDAVKGPGYDISRKY HDLEVNINLNGSGVYVYCFNFKIPEYVAVKSVYQMKFKAGNSDTEFLHAMNRVQGE EL"				
BASE COUNT	102 a	47 c	68 g	96 t	
ORIGIN					
Alignment Scores:					
Pred. No.:	6.95e-09	Length:	313		
Score:	262.50	Matches:	51		
Percent Similarity:	65.09%	Conservative:	18		
Best Local Similarity:	48.11%	Mismatches:	30		
Query Match:	6.65%	Indels:	7		
DB:	1	Gaps:	2		
US-09-494-297-2 (1-757) x AF009913 (1-313)					
0Y	51 PheGlyLeuVal-----GlusSerSerThrProAsnAlaIleAsnProasp	65			
			:::	:::	
Db	2 TTTGGTCAAGTAGGCTATTCTGCGGATGAGAACACTGTGCCGAATTTTAAAGCCAGAT	61			
0Y	66 SerSerSerGluTyrAgtTPTTPTTgLyGlyGluSerTyrValalagGlyHISProTyrTyr	85			
	:::		:::	:::	
Db	62 -----CCAGTWTTCCTGGTATGAGTATGATGATGTAAGGGGTATCCTGGATAT	115			
0Y	86 LysGlnPheArgValAlaHISAspLeuArgValAsnLeuGluGlySerArgSerTyrGln	105			
	:::				
Db	116 GATATCAGTAGTAGTATTATCATGATTATTAAGGTAAATCTTAAACGAGCTACAGCTACCA	175			
0Y	106 ValTyrCysPheAsnLeuLysLysAlaPheProLeuGlySerAspSerSerValLysLys	125			
Db	176 GTTATTATGCTTATTAATTCAAAAATAATTTTCCGACAAATGTGAATAACGTTACTACAAA	235			
0Y	126 TrpTyrLysLysHISAspGlyIleSerThrLysPheGluAspTyrAlaMetSerProArg	145			
Db	236 TGGTTTAAAAAGCCGGAGGAATTTCGATACCTTTCGACTTATGCATGAATGAATCCAGA	295			

```

OY      146 1lethrglyaspgluleu 151
      ::::::::::::::
Db      296 GTCCAGGAGAAAGAACTT 313

RESULT 14
AF009910      326 bp  DNA      linear  BCT 13-FEB-1998
LOCUS      Streptococcus pyogenes fibronectin binding protein F gene, partial
DEFINITION  cds.
ACCESSION  AF009910
VERSION    AF009910.1  GI:2267173
KEYWORDS
SOURCE     Streptococcus pyogenes
            Streptococcus pyogenes
ORGANISM   Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
            Streptococcus.
REFERENCE  1 (bases 1 to 326)
            Katerov,V., Andreev,A., Schalen,C. and Totolian,A.A.
            Protein F, a fibronectin-binding protein of Streptococcus pyogenes,
            also binds human fibrinogen: Isolation of the protein and mapping
            of the binding region
JOURNAL    Microbiology 144 (Pt 1), 119-126 (1998)
MEDLINE    98129085
PUBMED     9467904
REFERENCE  2 (bases 1 to 326)
            Katerov,V.E., Andreev,A.S., Schalen,C. and Totolian,A.A.
            Direct Submission
AUTHORS    Submitted (23-JUN-1997) Medical Microbiology, Lund University,
            Solvegatan 23, Lund 22362, Sweden
FEATURES
            source
            1..326
            /organism="Streptococcus pyogenes"
            /mol_type="genomic DNA"
            /strain="M4 strain M4/118"
            /db_xref="taxon:1314"
            <1..>326
            /codon_start=2
            /transl_table=11
            /product="fibronectin binding protein F"
            /protein_id="AAC38157.1"
            /db_xref="GI:2267174"
            /translation="FGQVAVAADEKTVPNRINSNPEPMGYDYSGRLLRYHNKLVN
            LNSKEYOAYCFNLKRFEPKKESSSPMWYKKLKDOSTFTFKYAEINPFGBELRHII
            LKVLVM"
BASE COUNT  118 a      41 c      75 g      92 t
ORIGIN
Alignment Scores:
Pred. No.:      1.62e-07      Length:      326
Score:          241.50      Matches:      51
Percent Similarity: 59.13%      Conservative: 17
Best Local Similarity: 44.35%      Mismatches: 34
Query Match:    6.12%      Indels:      13
DB:            1      Gaps:      3

US-09-494-297-2 (1-757) x AF009910 (1-326)
OY      51 Pheglyleuval-----GluSerSerThrProAsnAlaIleAsnProasp 65
      ||||| |||
Db      2 TTTGGTCAAGTAGCCTATGACGCGATGAGAAAGCTGCGCAATCGTATTAAT----- 55
      |||||
OY      66 SerSerSerGluTyrArgTPTPTGlyTyrGluSerTyrValArgGlyHisProTyrTyr 85
      |||||
Db      56 TCAATVCTGAGTTCTCTGGTATGCTTATGATTGCTATACCGAAGATTGTTAAGATAT 115
      |||||
OY      86 LysGlnPheArgValAlaHisAspLeuArgValAsnLeuGluGlySerArgSerTyrGln 105
      |||||
Db      116 -----CACATCTGAAGTAATATCTAAATGGAAGTAGAGATATCAA 157
      |||||
OY      106 ValTyrCysPheAsnLeuLysLysAlaIlePheProLeuGlySerAspSerValLysLys 125
      |||||
Db      158 GCGTATGTGTTTAACCTGAAAGGTTTGAGCCAAAGAAAGAAAGTAGAGAGTCTCTAAT 217

```

```

OY      126 TrpTyrLysHisAspGlyIleSerThrLysPheGluAspTyrAlaIleMetSerProArg 145
      |||||
Db      218 TGGTATTAATAAATTGATGCGAGTACAGAACTTTTAAATAATACGCGAAACCTAGG 277
      |||||
OY      146 1lethrglyaspgluleuasnGlnLysIleuArgValaIleMetTyr 160
      ::::::::::::::
Db      278 TTTACTGGAGAGAGATTAAGACGACATATATGAGAGGTCTCTAC 322

RESULT 15
AE009964      11264 bp  DNA      linear  BCT 03-APR-2002
LOCUS      Streptococcus pyogenes strain MGAS8232, section 12 of 173 of the
DEFINITION  complete genome.
ACCESSION  AE009964 AE009964
VERSION    AE009964.1  GI:19747406
KEYWORDS
SOURCE     Streptococcus pyogenes MGAS8232
            Streptococcus pyogenes MGAS8232
ORGANISM   Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
            Streptococcus.
REFERENCE  1 (bases 1 to 11264)
            Smoot,J.C., Barbican,K.D., Van Gompel,J.J., Smoot,L.M.,
            Chaussee,M.S., Sylva,G.L., Sturdevant,D.E., Ricklefs,S.M.,
            Porcella,S.F., Parkins,L.D., Beres,S.B., Campbell,D.S., Smith,T.M.,
            Zhang,Q., Kapur,V., Daly,J.A., Veasy,L.G. and Musser,J.M.
            Genome sequence and comparative microarray analysis of serotype M18
            group A Streptococcus strains associated with acute rheumatic fever
            outbreaks
JOURNAL    Proc. Natl. Acad. Sci. U.S.A. 99 (7), 4668-4673 (2002)
MEDLINE    11917108
PUBMED     21927593
REFERENCE  2 (bases 1 to 11264)
            Smoot,J.C., Barbican,K.D., Van Gompel,J.J., Smoot,L.M.,
            Chaussee,M.S., Sylva,G.L., Sturdevant,D.E., Ricklefs,S.M.,
            Porcella,S.F., Parkins,L.D., Beres,S.B., Campbell,D.S., Smith,T.M.,
            Zhang,Q., Kapur,V., Daly,J.A., Veasy,L.G. and Musser,J.M.
            Direct Submission
AUTHORS    Submitted (31-JAN-2002) Laboratory of Human Bacterial
            Pathogenesis/ROCKY Mountain Laboratories/NIH, 903 S. 4th St.,
            Hamilton, MT 59840, USA
FEATURES
            source
            1..11264
            /location="Qualifiers"
            /organism="Streptococcus pyogenes MGAS8232"
            /mol_type="genomic DNA"
            /strain="MGAS8232"
            /db_xref="taxon:186103"
            223..2313
            /gene="prtF21"
            /note="best blastp match p1r1730856 protein F2 -
            Streptococcus pyogenes"
            /codon_start=1
            /transl_table=11
            /product="protein F2-like protein"
            /protein_id="AAL96942.1"
            /db_xref="GI:19747407"
            /translation="MTQKNSYKLSFLSLTGTILGLILVFLIGLSGVSHAEETRNGAN
            KQAGEIKKNSQSEYENYEVDNRNIILODGEKLEIKRWDGQYOGFCQOLTKNFP
            TAGQVSKRLTKLSSDEETLKQVASKTSNRGTSGLNKLKQALVLEGGPTPKSD
            MNLGITEKLEIVODATIVYFETTVPRADSTRNNVNSQKKEVYQKLIIDTDIDKY
            EDVQDLDFVPODTNLQAVISVPEVLESILPWTSLKPIAOKDITPAKTIWDAREKPIIT
            FKLKQLDGEKEVAADAELKQINSEGOEISVYTNOLVDEKMAVIVSEKVDN
            GELLEPKDIKIKEDGLVTNTYVYKPTSGHYDEYFGNGHIDITEEDTPTDIYSGENQ
            KOIEEDSKPIDEVENNLIEGKNTMGEGDGTGNSKNEVEEDSRPYDPTLSGSSSD
            GSGDMTEEDSATPIKRSKDIQKELAGATMELRDSGSGTIYSWISDGGVQKDYLM
            PKRTFVETAPADGCEIATATFTPTNDEGGQYTVNKAIRKGAHITWVDAHYPTKPSGO
            VIDIEKLPDQGHSGSTTELEDKSSDVYIIGGQQIVETTEDTGTGHWGSGCTEY
            EDTKLVOSEFHDNKESESNSBIPKKDKPSMTSLSPATEKQHNMFPMVVTSCSLISSV
            FVISLTKTKRLSSC"
            complement(2579..3244)
            /gene="spyM18_0133"

```

```

gene
CDS
    complement(2579..3244)
    /gene="spym18_0133"
    /note="best blastp match gb|AAK33244.1| (AE006483)
    hypothetical protein [Streptococcus pyogenes M1 GAS]"
    /codon_start=1
    /transl_table=11
    /product="hypothetical protein"
    /protein_id="AAL96943.1"
    /db_xref="GI:19747408"
    /translation="MKRCNKLTFTSLAASMLSSYSKSVAHNDIKKGVHLWOCN
    IYVDNIVDTIKTLVDNNNSLEMTVTENSAKQWVYPDSVFLPKGVAKPEIYEHHY
    WDTGVHSKRSNTQMHYDMENQKTNFNOEDKPEPYGMSPSLDKFKLNDGFSFV
    LVDYGRQSHSTYFSHKWKWKFTELEGNKDKMKLPEFIAGIKONMPLAASPSYKGE
    FGE"
    3597..5003
    /gene="spym18_0134"
    3597..5003
    /gene="spym18_0134"
    /note="best blastp match gb|AAK33245.1| (AE006483)
    putative short-chain fatty acids transporter
    [Streptococcus pyogenes M1 GAS]"
    /codon_start=1
    /transl_table=11
    /product="putative short-chain fatty acids transporter"
    /protein_id="AAL96944.1"
    /db_xref="GI:19747409"
    /translation="MNNIKTKRRPMDRYIDGFMKWPESLFIQILFELVTVMSVMT
    DSPSIFGEKGGIIGVNGFNGKELISFAMOMTLLATGNVASSPRAHKPKFLAKP
    QTRQIFISIVGSIQFGLHKGIGMNAVYFGKELVQARQKIKVHTLFAVTLF
    TFLPATISLGAVALVSATPDLKNSVADAVKOVPSVPLTESVNLPLISLVCM
    LVPFCFALLAHPRDEKIMELDEIYHNSLDTASHVYIARTPAEKMANARLVIVG
    GAIVSYSLHRSVVGSLGDLNLSFNELFLGILLCGOGEPEYGSLEKGVNWSML
    VLOPEFYAGIFGIQSTGLGLEISHFEVAISNGTVPFVAYLISALINIVPSGGSF
    VIAPRYVPATIEGNDLGIQLAQYGDAITNLIYPMFALSYLSNKKLFNQVAT
    IPCVIVTGAIIYLYEY"
    complement(5073..5984)
    /gene="spym18_0135"
    complement(5073..5984)
    /gene="spym18_0135"
    /note="best blastp match gb|AAK34404.1| (AE006595)
    putative transcriptional regulator (lysr family)
    [Streptococcus pyogenes M1 GAS]"
    /codon_start=1
    /transl_table=11
    /product="putative transcriptional regulator"
    /protein_id="AAL96945.1"
    /db_xref="GI:19747410"
    /translation="MNISLAFIEIYVNCDFNLTKAQKSHISOPALSSYIKITEKEE
    GIVLFQRQKRLVGLTSVGERFENAKVIVASHQQLMADLRSSKRYKGEISIGIPOL
    ISGIVSELIPKLIENPEVKPHIVERGAYELQKFLKLEVDFAILLPTLTLLTAD
    IDILLRDLCLVMSHKTFOETPDSPNPMISIKALHROMLVTFDMFMINHQLTFEEA
    NLSLPVVCYASWDFLLNCTRSSYLLIPLRPITTHFAMMDGLVEVQLTEHPKKEVYIA
    SLKHNTKSHLKHVIKHTIIDYFTNQER"
    6105..7289
    /gene="spym18_0136"
    6105..7289
    /gene="spym18_0136"
    /note="best blastp match gb|AAK33246.1| (AE006483)
    putative acetyl-CoA acetyltransferase [Streptococcus
    pyogenes M1 GAS]"
    /codon_start=1
    /transl_table=11
    /product="putative acetyl-CoA acetyltransferase"
    /protein_id="AAL96946.1"
    /db_xref="GI:19747411"
    /translation="MTRKVITSAVPTPIQNFQGVKESLSAVDLGVTVTYKLIADTCL
    KSDAIDEVIFGNVLHAGLGQVAVRQVALNGLSDYDPATFIDVCGSGSLAVALGAK
    IQTGMADIVLVGGTENNSSQAPVYLOGQRWMSRSDSVYDTMKDGSADAFAGYAGAK
    TAENIVQOYGLTREEDQAFADSORKQALIEKGRFEETAPVTIIPORKEPPLIIDD
    EYKPGFTTVYKLAKRPAPFKIDSGTVAGASGIGNGAALILMSKREKAEELGIPILA
    KITSYASACVDSPTMCGPITPATKKALAKQULITDDILEANETAFRAQOLAVSRDGL
    FDNKRVNNGGATLGHPIGASGARILITVLLAEAKRDVHRGLATLCTIGGGQOSITV
    TR"

```

```

gene
CDS
    7301..7960
    /gene="spym18_0137"
    7301..7960
    /gene="spym18_0137"
    /note="best blastp match gb|AAK33247.1| (AE006483)
    putative acetyl-CoA:acetoacetyl-CoA transferase alpha
    subunit [Streptococcus pyogenes M1 GAS]"
    /codon_start=1
    /transl_table=11
    /product="putative acetyl-CoA:acetoacetyl-CoA transferase
    alpha subunit"
    /protein_id="AAL96947.1"
    /db_xref="GI:19747412"
    /translation="MKENKRIATAEAIISHTKDDDTIIVGGGMAATPEALIDALVDKG
    TKDLTLCNDAGFVDRCVGMKMANHOFKTYIAHGIKNEAGROMTACETTELIFOG
    TPARKIRIGATIGGRTPIGVGLVAEGKETIKTKITLLETPEDADVALIFAKA
    DEMGNLQYSGSENNFQDMAACAKTIVQAREIVPGTIOPEYVHPHIFVDIVKEG
    V"
    7963..8610
    /gene="spym18_0138"
    7963..8610
    /gene="spym18_0138"
    /note="best blastp match gb|AAK33248.1| (AE006483)
    butyrate:acetoacetate CoA:transferase, beta subunit
    [Streptococcus pyogenes M1 GAS]"
    /codon_start=1
    /transl_table=11
    /product="butyrate:acetoacetate CoA:transferase, beta
    subunit"
    /protein_id="AAL96948.1"
    /db_xref="GI:19747413"
    /translation="MAVALSKSEIQNRKARVAKLEDGLTVNLGIGLPTKVANFPE
    EMVYVPOSENGKSDPSDPSITVYAGGQVYVYVPGAFNSDPSFQIINGGVYD
    LTVIAGLEIENGDINXYLIPGMVYVCGMGAMPLIGAKKVIYAMHTNNGKRLIKE
    CTRPURAQVVDIITIMGVYKPTPDGIGQVIEISEGTFPEVQAATGVPILSIEG"
    complement(8732..9412)
    /gene="spym18_0140"
    complement(8732..9412)
    /gene="spym18_0140"
    /note="best blastp match gb|AAK33249.1| (AE006483)
    conserved hypothetical protein [Streptococcus pyogenes M1
    GAS]"
    /codon_start=1
    /transl_table=11
    /product="conserved hypothetical protein"
    /protein_id="AAL96949.1"
    /db_xref="GI:19747414"
    /translation="MDKETLNTWKVTVITFLHDVLDGNYEITLHVLDKNDIYIGELVNS
    HIGRSKOSPLETEFADLTINKVYKEDVTVNYKAVSPQAKVRSSTFPIKDKGNL
    EGMICIMDISAOGVARDLKLVINLLEHFIPTAKVPKTVTPQPEAVVAILTSNIOD
    IIGQIDPSSILRRNHVLSQVVKIDIVAKIYKGVFQLKGVANSKAVDILCISEBSVYRY
    LKKRIADDO"
    9586..9951
    /gene="spym18_0141"
    9586..9951
    /gene="spym18_0141"
    /note="best blastp match gb|AAK33250.1| (AE006484)
    conserved hypothetical protein [Streptococcus pyogenes M1
    GAS]"
    /codon_start=1
    /transl_table=11
    /product="conserved hypothetical protein"
    /protein_id="AAL96950.1"
    /db_xref="GI:19747415"

```

Alignment Scores:

Pred. No.:	0.000104	Length:	11364
Score:	224.50	Matches:	167
Percent Similarity:	35.69%	Conservative:	101
Best Local Similarity:	22.24%	Mismatches:	254
Query Match:	5.69%	Indels:	229
DB:	1	Gaps:	43


```

QY      620 -----ThrLysAspPheHisPheGluIleGluLeuLysAsnAsnLysGln 634
Db      2095 GAAGATACCTAACTAGTACAAATCCTCCACTTTGAT-----AACCAAGGAATCA 2142
QY      635 GluLeuLeuSerGlnThrValLysThrAspLys 645
Db      2143 GAAAGTAACTCTGAGATTCCTAAAGAAATAG 2175

```

Search completed: August 19, 2003, 17:28:07
 Job time : 6570 secs